HERC's Average Cost Datasets for VA Inpatient Care 1998 - 2005

Todd H. Wagner, PhD
Shuo Chen, PhD
Wei Yu, PhD
Paul G. Barnett, PhD

July 2006

Acknowledgments

A number of individuals have made this research possible and to them we owe our thanks. First, we would like to acknowledge Douglas Bradham, Ann Hendricks, Denise Hynes, Terri Menke, Anne Sales. Their insight, experience and advice helped to produce a higher quality product. We would also like to acknowledge Health Services Research and Development Service (HSR&D) and the Cooperative Studies Program (CSP) for funding this research.

Citing this handbook

Wagner TH, Chen S, Yu W, Barnett PG. “HERC’s Average Cost Datasets for VA Inpatient Care: 1998-2004.” Health Economics Resource Center, Department of Veterans Affairs, Menlo Park, California, 2005.

Note to Readers

In 2003, we published a supplement in Medical Care Research and Review that included papers based in part on the work presented in this manual. Readers are encouraged to cite those papers as the definitive source in future research articles. Copies of the articles are available on our website and upon request. The articles include:

Barnett, P. G., and Wagner, T. H. “Department of Veterans Affairs (VA) operates one of the largest integrated health care systems in the United States. Preface,” *Med. Care Res. Rev.* 60 (2003) 7S-14S.

Wagner, T. H., Chen, S., and Barnett, P. G. “Using average cost methods to estimate encounter-level costs for medical-surgical stays in the VA,” *Med. Care Res. Rev.* 60 (2003) 15S-36S.

Yu, W., Wagner, T. H., Chen, S., and Barnett, P. G. “Average cost of VA rehabilitation, mental health, and long-term hospital stays,” *Med. Care Res. Rev.* 60 (2003) 40S-53S.

Phibbs, C. S., Bhandari, A., Yu, W., and Barnett, P. G. “Estimating the costs of VA ambulatory care,” *Med. Care Res. Rev.* 60 (2003) 54S-73S.

Table of Contents

Tables

Figures

Chapter 1. Overview

For more information see, Barnett, P. G., and Wagner, T. H. "Preface," *Med. Care Res. Rev.* 60 (2003) 7S-14S.

The U.S. Department of Veterans Affairs (VA) provides health care to veterans at more than 120 inpatient facilities. Abstracts of all inpatient utilization are centrally available at the Austin Automation Center. However, no encounter-level charge or cost information is present. This is because the VA does not routinely generate patient bills. Consequently, VA researchers have not had economic data to estimate the cost of health care encounters.

In 1999, the VA funded the Health Economics Resource Center (HERC) to adapt existing cost methodologies (Barnett PG, Chen S, and Wagner TH, 2000) and to expand methods where possible and necessary. The current methodology, described in detail in this manual, is evolving and continues to improve over time. Input from users is crucial so that improvements can be made. We welcome all suggestions.

This report describes HERC's method for estimating the cost of VA inpatient stays in fiscal years 1998-2003.[1] Our goal was to develop a set of long-term costs that could be used in cost-effectiveness analysis. We use the term long-term in the economic sense that all costs are variable. A companion report on outpatient costs is also available on our web site (http://www.herc.research.med.va.gov/publications/default.asp).

Known as the "average cost" method, we assume that every health care encounter has the average cost of all encounters that share its same characteristics. Although this assumption limits the accuracy of the cost estimates, especially for outliers, this is the only available method of generating a comprehensive set of encounter-level estimates of all patient care provided by VA. The average cost method relied on the following assumptions:

- To find the cost of rehabilitation, blind rehabilitation, spinal cord injury, psychiatric, substance abuse, intermediate medicine, domiciliary, and psychosocial residential rehabilitation stays, we found the average cost of a day of stay, and multiplied it by length of stay to estimate the cost of care. This makes the assumption that every day of stay has the same cost, that is, that costs are directly proportionate to the length of stay. This type of care is hereafter referred to as non-medical/surgical or rehabilitation, mental health or long-term care.

---

[1] The federal fiscal year begins on October 1 and ends on September 30 of the following year. The convention is to refer to a federal fiscal year (FY) by the year it ends, thus FY98 represents the period October 1, 1997 to September 30, 1998.

- To find the cost of acute medical-surgical hospital care, we built a cost function using relative value units (RVUs) from the non-VA sector. These RVUs were the Diagnosis Related Group (DRG) weights used by the Centers for Medicare and Medicaid Services (CMS) to reimburse U.S. hospitals for the care they provide to Medicare patients. The RVUs reflect the effect of diagnosis on the relative quantity of resources used in a hospital stay. In addition to DRG weights, the cost function included length of stay, demographic and other clinical information. The method we employed makes the following assumptions: (1) that the non-VA relative value units, the Medicare DRG weights, reflect the relative costs of VA hospital stays, and (2) that all stays with the same characteristics have the same cost.

- To find the cost of long-term care, we employed relative value weights known as resource utilization groups (RUG). Therefore, costs of long-term care are adjusted for case-mix as measured by the RUG score. Veterans with higher RUG scores are considered to have higher costs (FY98-00 only).

- In FY01 - FY04, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. A preliminary review of these new RUG data suggests that ongoing data monitoring is needed before they can be used to determine costs.

1.1 Changes in methods over time

As the average cost method evolves, improvements are made. Below is a brief summary of the changes that were adopted with the FY98-FY04 datasets.

1.1.1 Acute medical-surgical short stay hospitalizations

Beginning in FY98 we used a cost function based on Medicare data. We made the cost function's form highly flexible to account for variations in severity and length of stay.

1.1.2 Categories of inpatient care

From FY98 - present, we estimate inpatient costs for eleven categories of care.

1.1.3 Nursing home care

For FY98-FY00, the Resource Utilization Groups (RUGs) were used to case-mix adjust the average daily cost. This is discussed in greater detail in Chapter 6. In FY01, the cost of long-term care is a per diem rate because VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are available, but additional monitoring is needed before the data can be used to estimate costs.

1.1.4 DSS Costs

In FY04, we switch from using the Cost Distribution Report (CDR) to a department- level summary from the Decision Support System Nation Data Extracts.

1.2 Organization of User Guide

The average cost estimates represent a merger of centralized VA cost and utilization databases and relative value units obtained from non-VA databases. This paper begins with a description of the VA Cost Distribution Report (CDR), our source of VA cost information. Section 3 covers the utilization data. Chapter 4 provides an overview of our method of merging the CDR with the VA utilization files.

Section 5 describes our method of determining the daily cost of non-medical/surgical care: rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, domiciliary, and intermediate medicine. Chapter 6 describes the methods for estimating the cost of nursing home stays. Chapter 7 describes our method of finding the cost of acute medical-surgical hospital stays. Chapter 8 is the user's guide.

Chapter 2. Cost Data

We used the Cost Distribution Report (CDR) to create the 1998-2003 HERC average cost datasets. The CDR ceased production in 2004. So for the 2004 HERC average cost datasets we summarized the DSS National Data Extract (treatment specialty file) to create department-level costs. We then used this summary to estimate the encounter-level HERC costs. Below is a brief discussion of the CDR and DSS NDE.

2.1 Cost Distribution Report

The Cost Distribution Report (CDR), also called report RCS 10-0141, is routinely prepared by all VA medical centers. The CDR represents an estimate of the costs expended by each VA patient care department.

VA expenditures are recorded in its general ledger, the Financial Management System (FMS). The FMS system tracks expenditures by cost center, a budget entity that corresponds to a VA service. Examples of VA cost-centers are Medical Service, Nursing Service, and Plant Operations. Cost centers do not correspond to a specific patient care department.

The CDR is created by distributing costs reported in the FMS cost centers to the "cost distribution accounts" (CDA) of the CDR. The CDAs include patient care departments, such as Medical Intensive Care, or Ambulatory Care, Medicine. CDAs also include indirect cost departments.

The distribution of costs is based on estimates prepared by the service chiefs in each medical center. Each service chief estimates the amount of time staff spent on different activities. The cost of staff time, as reported in FMS, is then assigned to each CDA. At the end of each fiscal year, a cumulative CDR is prepared, and it is reconciled to the costs reported in FMS.

Our average cost estimate required information about the cost of each category of inpatient care, including its share of indirect costs. The CDR distributes indirect costs only to groups of patient care departments. We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs.

2.2 CDR units and unit costs

We did not use the units of service or the unit costs reported in the CDR because of our concerns in the accuracy of these data. Rather than use these units or unit costs, we used the VA Patient Treatment Files as our source of utilization data to find unit costs (see Chapter 3).

2.3 DSS Summary

The CDR ceased production in 2004. For a department-level cost dataset, we chose to create our own from the DSS National Data Extract Treatment Specialty File (TRT). The TRT is a encounter-level dataset that tracks the bedsection (what they call the treatment specialty). The bedsection number is identical to the treatment specialty, just a change in names. By summarizing the TRT into a department level dataset, we ensure that the HERC and DSS NDE's are based on the same underlying costs. In prior years this was not the case. For example, prior to FY04, HERC excluded central office costs. Therefore, when researchers compared HERC and DSS costs from FY03, the datasets differed in both the underlying costs and the relative value units. Now, with FY04 the underlying costs are the same and the only difference between the datasets is the relative value units.

Table 3.1 shows the DSS TRT and CDR costs for 2003. The total costs are approximately $600 million (5.4%) apart, with the CDR reporting more inpatient costs. However, this small difference masks much larger differences by categories. Surgery, internal medicine, nursing home and PRRTP have differences in excess of 20%. Also CDR costs were higher in every category but domiciliary, nursing home and PRRTP.

Table 3.1: DSS and CDR Cost Comparison for FY03

	DSS Total Costs		**CDR Total Costs**		**National Cost**
	National	**Local**	**National**	**Local**	**Diff. (%)**
Medicine	3,132,024,993	3,133,360,986	3,329,463,481	3,329,315,342	5.9%
Rehabilitation	89,462,191	89,859,283	102,875,347	102,874,029	13.0%
Spinal Cord Injury	56,991,112	57,009,870	65,181,269	65,185,781	12.6%
Blind Rehab.	274,589,433	274,589,430	301,278,483	301,273,138	8.9%
Surgery	1,824,835,695	1,825,106,490	2,391,960,548	2,391,242,870	23.7%
Psychiatry	965,560,239	967,362,246	1,112,991,416	1,129,023,967	13.2%
Substance use	60,830,365	59,028,354	72,148,503	72,149,146	15.7%
Inter. Medicine	195,967,740	194,664,564	276,969,168	277,495,664	29.2%
Domiciliary	367,383,921	365,727,926	306,421,042	306,594,343	-19.9%
Nursing Home	2,194,938,443	2,195,875,004	1,756,350,731	1,759,736,090	-25.0%
PRRTP	112,262,178	112,262,178	87,082,592	87,082,356	-28.9%
Total	9,274,846,310	9,274,846,331	9,802,722,580	9,821,972,726	5.4%

PRRTP is psychosocial residential rehabilitation programs

The differences shown in Table 3.1 also exist in the calculation of the average per diem costs. Table 3.2 shows the per diem amounts for FY03.

Table 3.2: Average Costs Per Day for DSS and CDR in FY03

		DSS Average per diem costs		CDR Average per diem costs	
		National	Local	National	Local
Medicine					
	Mean	1,516	1,521	1,612	1,653
	SD	--	371	--	437
Rehabilitation					
	Mean	1,146	1,200	1,318	1,597
	SD	--	246	--	1,223
Spinal Cord Injury					
	Mean	789	765	902	890
	SD	--	182	--	215
Blind Rehabilitation					
	Mean	974	1,023	1,069	1,185
	SD	--	257	--	515
Surgery					
	Mean	2,412	2,285	3,163	3,547
	SD	--	534	--	1,880
Psychiatry					
	Mean	732	869	836	1,136
	SD	--	322	--	1,080
Substance use					
	Mean	540	4,871	641	7,339
	SD	--	20,335	--	30,910
Intermediate Medicine					
	Mean	908	1,003	1,284	35,830
	SD	--	502	--	235,845
Domiciliary					
	Mean	196	234	164	192
	SD	--	97	--	76
Nursing Home					
	Mean	475	503	380	413
	SD	--	129	--	143
PRRTP					
	Mean	305	330	237	232
	SD	--	314	--	139

PRRTP is psychosocial residential rehabilitation programs
SD denotes standard deviation; we report the standard deviations across medical centers

Chapter Summary

- HERC used the CDR for FY98-FY03 Average Cost Datasets.

- We created a department-level dataset from the DSS Treatment Specialty File to create the FY04-FY05 Average Cost datasets.

- The CDR was created by distributing costs reported in the FMS cost centers to the "cost distribution accounts" (CDA) of the CDR. The CDAs include patient care departments, such as Medical Intensive Care, or Ambulatory Care, Medicine.

- The DSS Treatment Specialty File is an encounter-level dataset, including direct and indirect costs. We summarized this file by bedsection (also known as treating specialty).

- We use the VA Patient Treatment File as our source of utilization data to find unit costs.

Chapter 3. VA Inpatient Databases

The VA maintains a database of hospital stays called the Patient Treatment File (PTF). Although this database contains neither cost nor charge data, it includes data such as patient demographics, length of stay, and the Diagnosis Related Group (DRG) for the hospitalization.

3.1 VA Utilization Datasets

The PTF records information on hospital stays in different files. It is important to understand how this information is organized because VA defines a hospital stay somewhat differently than non-VA hospitals.

There are three file types of files: observation, extended care and other care. The observation, extended care and other care have a main and a bedsection, and for each of these there is a discharge and census file. As shown in Figure 3.1, there are 12 files.

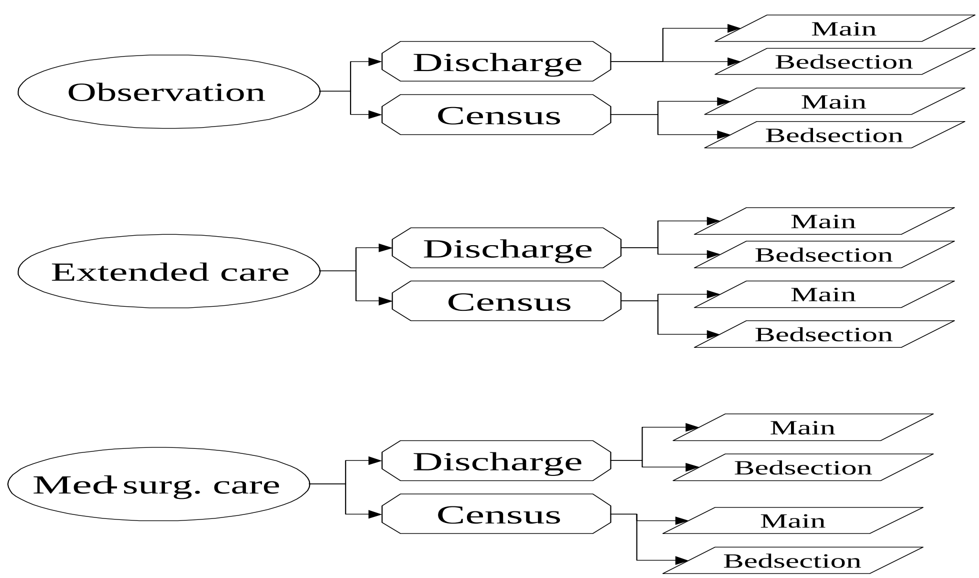


Figure 3.1: VA inpatient data files

3.1.1 PTF med-surg main discharge file (PTF Main)

This file reports all hospital stays that ended in a particular year. This file contains one record for each hospital stay. The main file does not use a definition of a hospital stay that is strictly comparable to non-VA hospitals. In the non-VA sector, an acute medical-surgical hospitalization followed by a long-term care stay would be recorded as two different stays. In the PTF main file, however, this is often recorded as a single stay.

In some cases, the PTF main is analogous to the non-VA sector. For example, an acute medical-surgical hospital stay that began in the Intensive Care Unit and ended in a medicine ward would be regarded as a single stay in the non-VA sector. This would be recorded as a single record in the PTF main file.

We wanted to apply relative value units from acute medical-surgical stays in non-VA hospitals to estimate the cost of acute medical-surgical VA hospital stays. This required us to develop a definition of what is an acute medical-surgical hospital stay. We used information from both the main and bed-section files to define an acute medical-surgical inpatient stay; see Chapter 7 for a description of our methods for finding the cost of acute medical-surgical hospital stays.

3.1.2 PTF acute care bedsection discharge file (PTF bedsection)

The PTF Bedsection file, is similar to the PTF Main, except that it has multiple records per stay. The PTF Bedsection file divides hospital stays into sequential segments, with one record for each portion of the stay spent in a different bedsection. A bedsection is a hospital ward such as medicine, intensive care, rehabilitation, or long-term care. The bedsection view provides information on the number of days the patient spent in each bedsection.

The PTF Bedsection file does not contain the same data elements as the PTF discharge main file. It is necessary to use both files to obtain all of the hospital discharge information that is required. For example, gender, age and number of diagnoses are in the PTF main discharge file but not in the bedsection discharge file.

There are other slight, but important distinctions between the PTF Main and Bedsection files. As mentioned above, the Main file does not use a definition of a hospital stay that is strictly comparable to the non-VA sector. Both acute medical-surgical stays and rehabilitation, mental health or long-term stays are all aggregated in the PTF Main, while the non-VA sector would typically have an acute medical-surgical stay record and a rehabilitation, mental health or long-term stay record. The PTF Bedsection file, on the other hand, separates stays into each bedsection stay. Hence a stay with an acute medical-surgical bedsection component and a rehabilitation, mental health or long-term stay would have two records, which is analogous to the non-VA sector. However, the PTF bedsection file also separates transfers between acute medical-surgical bedsection or between rehabilitation, mental health or long-term bedsections. Such transfers result in more than one record in the PTF Bedsection file. In the non-VA sector, transfers between acute medical-surgical wards would be considered as part of one stay as long as the patient was not transferred to a rehabilitation, mental health or long-term care ward during the stay.[2]

---

[2] The bedsection is the "treating specialty" assigned to the physician who is responsible for the patient's care. It roughly corresponds to the location where care is delivered. We used this variable from the PTF, called BEDSECN, to characterize inpatient care. PTF includes another

3.1.3 PTF Acute Census files

The PTF main and bedsection discharge files include information on all stays that ended during a given fiscal year, regardless of when they began. As is common with discharge files, they do not report on people occupying a bed at the end of the reporting period. To fill this gap, the PTF Census Files includes information on patients who are in a VA hospital at the end of the fiscal year. Note that Census files are given the name of that fiscal year in which they ended. For example, Census FY98 was completed in September 30 1998.[3]

3.1.4 PTF Extended care files

As shown in Figure 3.1, the inpatient utilization files at Austin are divided into three components pertaining to acute inpatient care, extended care, other observation stays. Most stays that start in a nursing home file are included in the extended care file, regardless of the bedsection in which the patients ends up. On the other hand, stays that do not start in the nursing home are usually listed in the non-extended care files, even if the patient was transferred to a long-term care unit.

Since stays may be made up of both acute medical-surgical and long-term care, both of these files contain information on stays that involve acute medical-surgical and long-term care bed-sections. The assignment of stays to one set of files or the other did not affect our treatment of data. We merely used all data from both sets of files for our calculations.

3.1.5 Observation Bed files

The Observation Bed file was first created in 1998 and has been used with increasing frequency in each year since then. If a stay includes an observation bedsection, then the observation portion of the stay is separated from the rest of the stay and included in this file. Most observation bed stays were one day stays, with the patient being discharged from the hospital. However, in some cases there are observation stays that preceded an acute medical-surgical hospital stay. In a few rare instances, people were discharged from an acute medical-surgical hospital stay to the observation bed. In the latter two examples, the portion of the stay that corresponds to the observation bed is kept in the observation bed file.

---

variable, PLBED, to denote the location where care was provided. We did not use this variable to characterize the location of care because many records have missing values for PLBED, whereas all records have a value for BEDSECN.

[3] The one cautionary note with the Census file is that not all bedsections are coded on September 30 or October 1. Some stays cross the fiscal year are logged in on October 2 and 3. On rare occasions, the stay may be logged in as late as two weeks after. To get an accurate estimate, rather than rely on Census counts for October 1, we recommend that people use the bedsection inday and outday variables to identify whether the person was in a bed at the end of the fiscal year.

Observation bedsections were created at the same time as the VA was implementing managerial performance incentives to reduce the number of inpatient days per 1000 treated veterans. Observation data are not included in this performance measure.

Because observation bed stays are so heterogeneous, they present some difficulty in determining their cost. We decided that all observation stays would be given the daily cost of the marginal cost of a day. To calculate the marginal cost of day, we used a statistical model with Medicare data (see Chapter 7).

Chapter summary

• The Patient Treatment File (PTF) records information on hospital stays are in two different datasets (PTF main and PTF bedsection).

• The PTF main file reports all hospital stays that ended in a particular year.

• The PTF utilization files (Main and Bedsection) are divided into three components pertaining to acute care, extended care, observation stays. Acute care, extended care and observation stays each have a discharge and a census file.

• The bedsection file divides hospital stays into sequential segments, with one record for each portion of the stay spent in a different bedsection. A bedsection is a hospital ward such as medicine, intensive care, rehabilitation, or long-term care.

• The PTF Main and Bedsection are discharge files and they do not report on people currently occupying a bed at the end of the fiscal year. To fill this gap, the PTF Census Files includes information on patients who are in a VA hospital at the end of the fiscal year.

Chapter 4. Merger of cost and utilization databases

This section describes how we merged the CDR with VA utilization databases. The VA database of hospital stays is called the Patient Treatment File (PTF). This paper does not cover outpatient data.

We excluded the cost of facilities that do not provide patient care. In addition, we accounted for mergers between VA medical centers. Over time, facilities have consolidated, but these consolidations were not necessarily implemented at the same time in the cost and utilization databases. We recoded data to keep a common definition of facility in the databases. Since patient care departments are sometimes defined differently in the cost data than in the utilization data, we aggregated departments to find a common denominator.

4.1 Excluded facilities

We excluded the 16 facilities that report costs in the CDR, but do not report utilization in either the PTF or the OPC. These include records for VA Headquarters (station 101), information services centers, and other VA support facilities. A list of these facilities, and their three digit facility number, is provided in Table 4.1. Nine of these facilities do not appear in the official listing of VA facilities.[4]

Table 4.1: Excluded Facilities

Facility Number	Facility Name
101	VHA Headquarters
200	Austin Automation Center
722	Albuquerque, NM Outpatient Center
741	Denver CHAMPVA
721, 724, 742, 760, 761, 762, 763, 764, 765	
792	Prosthetics Center
794	Somerville
797	Hines (CIO)

We felt that central administration may involve activities that are more characteristic of a health care payer, rather than a health care provider. For this reason, we decided not to count these facility's costs as overhead costs that should be distributed to patient care departments.

[4] Consolidated Address and Territorial Bulletin 1-L, U.S. Department of Veterans Affairs, Washington, DC 20420, August 31, 1999

4.2 Facility mergers

VA has been consolidating facilities. When one facility merges with another, they both take on a single identification number (see Table 4.2). This change is sometimes implemented at different times in the different data systems. We wished to maintain the distinction between facilities as long as it was possible. We also wished to work with observations that consisted of facility level data for an entire fiscal year. We consolidated all data into the new facility number in the first fiscal year that the CDR or the utilization databases no longer maintained a distinction between the facilities.

Table 4.2: Facility Consolidations in 1997-2005

	Old STA3N	New STA3N
1997		
VA Chicago Health Care System	535	537
VA Central Alabama Health Care System	680	619
VA North Texas Health Care System	522	549
Southern California System of Clinics	665,752	665
Hudson Valley VA Health Care System	533	620
VA Central Iowa Health Care System	592	555
VA Greater Nebraska Health Care System	574	597
1998		
VA Eastern Kansas Health Care System	686	677
VA Montana Health Care System	617	436
North Florida/South Georgia Veterans Health Care System	594	573
VA Greater Los Angeles Health Care System	752	691
1999		
Greater Los Angeles Health Care System	665	691
Boston VA Health Care System	525	523
2000		
NY Harbor Health Care System	527	630
Upstate NY Health Care System	532	528
Upstate NY Health Care System	670	528
VA Mid Tennessee Health Care System	622	626
Upstate NY Health Care System	500	528
VA Nebraska Western	584	636
2001		
Columbia MO Harry S Truman Memorial VA Medical Center	543	589
Eastern Kansas VA Health Care System	677	589
Marion IL VA Medical Center	609	657
Popular Blue MO John J Pershing Medical Center	647	657

2002		
VA Eastern CO Health Care System	567	554
Kansas City VA Medical Center	452	589
2003-2005		
None		

4.3 Definition of patient care unit

Patient care units are defined differently in the CDR and the utilization databases. In the CDR, care is characterized by the Cost Distribution Account. The Cost Distribution Report Handbook maps the correspondence between Cost Distribution Accounts and the utilization databases. It does not include the Cost Distribution Accounts and utilization codes created since 1996, so the handbook is now out of date.

The Patient Treatment File (PTF) characterizes inpatient care by bedsection, which refers to the ward where the patient received care, such as medical intensive care unit, or nursing home unit. Each inpatient Cost Distribution Account in the CDR reports the costs of operating a group of several different bedsections. To learn about the correspondence between new bedsection codes and new cost distribution accounts, we examined the variable BEDCDR in the PTF bedsection file. This variable has the value of the CDA that corresponds to the bedsection. Only one CDA is assigned to each bedsection. As a result, the exact correspondence between BEDCDR and BEDSECN (the variable for bedsection) in the PTF represents a statement of the CDA associated with each bedsection.

Our review of CDR data suggests that some medical centers do not consistently use the definitions given in the CDR handbook and in these supplemental sources. The cost of providing care in a particular bedsection is not always assigned to the corresponding CDA specified in the CDR handbook. Some facilities have utilization in bedsections without assigning any costs to the corresponding CDA. In other cases, costs are assigned to a CDA, but no utilization appears in the corresponding bedsections.

The cause of this problem is the addition of new CDAs to the CDR and new bedsections to the PTF. Facilities may implement new utilization codes and CDAs at different times.

We dealt with these potential problems by defining aggregate “patient care categories.” These categories represent our best judgment about what constitutes the smallest common denominator between cost and utilization. A patient care category represents a group of related cost distribution accounts, and their associated utilization.

We defined patient care categories based on earlier work (Barnett PG et al., 2000). We aggregated CDAs into eleven categories, and ascertained that for almost every medical center, if the category had costs, it also had utilization, and if it had utilization, it also had costs. We also examined the mean cost of care, examining outliers that suggested mismatch of cost and utilization data.

For some categories of care at some medical centers, there were mismatches between cost and utilization data. Most mismatches were handled by assigning the costs and utilization to a similar department, creating a higher level of data aggregation. For more details on the reconciliation, see 4.5.

4.4 Merger of cost and inpatient utilization data

The CDR reports on expenditures in a federal fiscal year, which runs from October 1 until September 30. As mentioned above, we wanted to identify the amount of care provided during the fiscal year. Since hospital stays may span fiscal years, we developed a method to divide hospital utilization between fiscal years.

The denominator for the cost data was the fiscal year, whereas the denominator for the utilization data was discharges. These denominators are not equivalent. We could have ignored this difference. This would have been equivalent to assuming that bed occupancy was constant over the year. However, this assumption would be wrong because we know that there is a trend to shorten length of stay and to reduce hospitalization. And we did not want to assume that the same number of patients is in the hospital at the start and at the end of the fiscal year.

A better way to adjust for the difference in denominators was to use information from the Census files. With the Census files we adjusted the discharge file so that it more closely approximated utilization in the fiscal year.

For the utilization data, we included days spent during the current fiscal year by all patients. For those discharged during the fiscal year, their data came from the PTF, limiting the days to those in the fiscal year. For those patients not discharged by the end of the fiscal year, we obtained these days of stay from the PTF census files. This calculation included "leave" days, that is, days that a patient was absent from a hospital, though not yet discharged. The PTF records leave days, but it does not indicate when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

The finest level of detail for the cost data is at department level; patient-level cost data do not exist. To merge the cost and utilization data, we identified 11 categories of inpatient care (see Table 4.3).

Table 4.3: Categories of Inpatient Care (excludes indirect costs)

Category of Care	CDR acct	BEDSECN
Inpatient Medicine	1110, 1114, 1117, 1118, 1119, 1120, 1130	1-12, 14-17, 18[#], 19, 24[#], 31, 34, 35, 75, 83
Inpatient Rehabilitation	1113	20, 41[#]
Inpatient Blind Rehabilitation	1115	21, 36[#]
Inpatient Spinal Cord	1116	22, 23[#]
Inpatient Surgery	1210-1213, 1230	50-63, 65[#]
Inpatient Psychiatry	1310, 1311, 1314, 1315, 1316, 1317, 1320, 1330, 1711[†], 1712[†], 1714[†], 1717[†]	25[†], 26[†], 28 , 33, 70, 71, 76, 77, 79, 89, 91, 92, 93, 94[#]
Inpatient Substance Abuse	1312, 1313, 1713[†], 1715[†]	27[†], 29[†], 72-74, 84, 90
Inpatient Intermediate	1610, 1620	32,40
Inpatient Domiciliary	1510, 1511, 1512, 1513^, 1520	37^, 85-88
Inpatient Long Term	1410, 1420, 1415, 1416, 1425	80, 81
PRRTP[†]	1711, 1712, 1713, 1714, 1715	25, 26, 27, 28, 29, 38[#], 39[#]

[†] These CDR accounts and bedsections were assigned to psychiatry and substance abuse at medical centers that did not have an official PRRTP program. In FY03 PRRTP programs existed at: 500, 501, 504, 463, 637, 515, 516, 518, 523, 532, 541, 549, 554, 561, 568, 573, 590, 459, 586, 589, 555, 595, 598, 546, 620, 622, 556, 631, 632, 635, 640, 645, 653, 658, 662, 663, 666, 656, 676, 678, 687, 689

^ New for FY00

[#] New for FY01

PRRTP programs are less intensive inpatient programs for psychiatry and substance abuse. A separate time series analysis confirms that medical centers that adopted PRRTP care had an associated decrease in the daily cost of substance abuse and psychiatric care, but this new program allowed them to provide more services and this offset the savings (Wagner & Chen, 2002).

4.5 Data reconciliation

After using the 11 inpatient categories to merge the cost and utilization data for each medical center, we reconciled the cost and utilization databases. This was necessary because the VA does not routinely reconcile these two databases. The most obvious discrepancies are when a category has costs but no utilization. The opposite can also be true– utilization exists without costs. In reality, the occurrence of these discrepancies is quite rare. When they occurred we merged the substance abuse costs and utilization with the psychiatry costs and utilization. Appendix A describes all the reconciliations that were done for FY98-present.

4.6 Daily rate

After reconciling the 11 inpatient categories, there was a direct correspondence between costs and utilization. We divided total costs by total utilization to find the average cost for each category of care at each medical center. We compared rates across medical centers, and we found the average rate for each of the categories. Table 4.3 lists the average rates for inpatient care in FY98 -present.

It is important to note that this daily rate does not account for case mix, clinical information or demographic characteristics. It is just an average daily rate. To use these rates, one would have to assume that costs are only a function of length of stay. This is not an extremely appealing assumption. Unfortunately, for most of the categories we have little additional information that can be used to make more accurate cost estimates. For acute medicine and surgery, we have a better method for estimating costs, which is covered in Chapter 7. For nursing home care, we have developed a new method that accounts for case-mix; this is presented in Chapter 6.

Table 4.4: Median facility cost per day of stay for inpatient care, FY98-04

	CDR Based Estimates						DSS
	FY98	FY99	FY00	FY01	FY02	FY03	FY04+
Inpatient Medicine*	$1,195	$1,304	$1,319	$1,381	$1,465	$1,600	$1,619
Inpatient Rehabilitation	$890	$1,029	$1,012	$1,102	$1,377	$1,318	$1,191
Inpatient Blind Rehabilitation	$728	$762	$815	$834	$861	$861	$903
Inpatient Spinal Cord	$764	$838	$791	$843	$971	$1,136	$1026
Inpatient Surgery*	$2,625	$2,797	$2,455	$2,700	$2,882	$3,190	$2,469
Inpatient Psychiatry	$680	$745	$744	$769	$864	$918	$861
Inpatient Substance Abuse	$821	$576	$418	$595	$666	$726	$643
Inpatient Intermediate	$625	$548	$525	$599	$794	$1,213	$990
Inpatient Domiciliary	$126	$238	$126	$162	$173	$168	$228
Inpatient Long Term	$275	$303	$305	$339	$358	$394	$536
PRRTP	$161	$179	$179	$213	$220	$239	$291

Costs presented are the median of each facilities average daily cost.
Includes overhead/indirect costs
* We do not recommend using this cost as we have provided more accurate estimates (see Chapter 7).

Chapter summary

- We excluded the cost of 16 facilities that do not provide patient care. We felt that central administration may involve activities that are more characteristic of a health care payer, rather than a health care provider.

- We also accounted for mergers between medical centers. If medical centers merged during a fiscal year, we merged their utilization and cost data for the entire fiscal year. It was not possible to separate costs and utilization before and after the merger.

- Patient care units are defined differently in the CDR and the utilization databases. In the CDR, care is characterized by the cost distribution account. The Patient Treatment File (PTF) characterizes inpatient care by the "bedsection."

- Our review of CDR data suggested that many medical centers do not consistently use the definitions given in the CDR handbook. We dealt with this by defining aggregate 11 "patient care categories."

- In merging the PTF data to the CDR data, one must remember that the PTF has a discharge view while the CDR takes a fiscal year view. These are not synonymous views and an adjustment is needed to make these equivalent.

- Even for these patient care categories there was not always a one to one correspondence between the CDR and the PTF. We did our own reconciliation to solve this problem. The exact reconciliations are provided in an Appendix A.

After reconciling the 11 inpatient care categories, we generated an average cost per day in each category.

Chapter 5. The cost of rehabilitation, mental health and long-term care

5.1 What is rehabilitation, mental health and long-term care?

Most US hospitals differentiate between short-stay acute medical-surgical and non-medical/surgical hospitalizations. Short-stay acute medical-surgical hospitalizations are generally for acute medicine and surgical treatment. While over 90% of short stay hospitalizations are less than 60 days long, there are rare cases that involve a length of stay up to and over a year. In the VA, about half of the inpatient stays can be categorized as acute medical-surgical defined by their bedsections (see Table 4.3). The remaining stays include rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, intermediate care, domiciliary, and nursing home. This chapter describes how we estimated the cost for rehabilitation, mental health or long-term care for FY98-FY03. The one difference between FY98 and prior years is the use of case-mix adjustment for nursing home care for FY98-FY00. After FY00, nursing home care is based on a per diem cost. More information on the cost of nursing home care is covered in Chapter 6.

5.2 Cost methodology for rehabilitation, mental health and long-term care

Determining costs for rehabilitation, mental health and long-term care is the most straightforward of the cost determination methods. The premise is to merge the CDR and PTF bedsection databases for each of the 11 care categories. The 11 care categories are defined by bedsection and cost distribution accounts (see Table 4.3). Two values are needed to calculate a daily cost for each of the care categories: total costs and total number of days. With this information, a daily rate can easily be calculated by dividing total costs by total days. This can be done either at the medical center level or for the entire nation. When this is done at the level of the medical center, the result is an average daily rate for that medical center. We refer to this rate as the local daily cost estimate.

5.2.1 Leave and pass days

For stays that began before the beginning of the fiscal year, we found the length of stay during the current fiscal year by finding the number of days between the discharge date and the beginning of the fiscal year. This calculation considered "leave" days, that is, days that the patient was absent from the hospital, though not yet discharged. Leave days are also called Absent Bed Occupant Days and are given the variable name LVB in the PTF. The PTF records leave days in a variable named LVB, but it does not record when they occurred. We assumed that leave days are uniformly distributed throughout the stay.

5.2.2 Local outlier costs

As one might expect, there is more variation in the local daily rates than the national daily rates. This raises the question about the accuracy of the local rate. To help identify inaccurate local costs, we generated a flag if a medical center had a daily rate that ± 2 standard deviations from the average of all VA medical centers (for that particular care category). Part of this variation

Chapter summary

- Almost half of inpatient VA stays are rehabilitation, mental health or long-term . We categorize non-medical/surgical care into nine categories: to rehabilitation, blind rehabilitation, spinal cord injury, psychiatry, substance abuse, intermediate care, domiciliary, and nursing home.

- Except for nursing home care, our cost methodology is to generate an average daily rate for each category.

- The average daily rate was estimated for each medical center, providing a local cost estimate, or at the national level, providing a national cost estimate.

- As one might expect, there is more variation in the local daily rates than the national daily rates.

We generated a flag if a medical center had a daily rate that was ±2 standard deviations from the average of all VA medical centers (for that particular care category). Part of this variation could be explained by factors such as wages. However, some of this variation is due to accounting mistakes or inconsistencies. Therefore, one should be informed and check for outliers when using the local cost estimates.

could be explained by factors such as wages. However, some of this variation is due to accounting mistakes or inconsistencies. Therefore, the flag variable allows the analyst to check for outliers when using the local cost estimates.

5.2.3 Why local rates at all?

Given that there is more variation in the local rates than the national rates, one may ask why do we calculate local rates at all. The answer is that sometimes the variation in the local rates is important. Wages are one factor that affects costs, as they depend on the labor market in different geographic localities. If a researcher is interested in the effect of an intervention on a local medical center or VISN, then the local rates may be more appropriate because they partly reflect the wage differentials and other local differences.

5.2.4 Adjusting for case-mix

Although DRGs have been created for mental health and rehabilitation stays, the cost of stays assigned to these DRGs is highly variable. Because DRGs do not explain the variation in cost of rehabilitation and mental health stays, facilities that provide this sort of care were exempted from the Prospective Payment System of Medicare. We estimated the cost of this type of care using the average daily cost.

Chapter 6. The cost of nursing home care

For more information, see: Yu, W., Wagner, T. H., Chen, S., and Barnett, P. G. “Average cost of VA rehabilitation, mental health, and long-term hospital stays,” *Med. Care Res. Rev.* 60 (2003) 40S-53S.

VA long-term care patients are evaluated using the Resource Utilization Group (RUG) assessment method. These assessments are performed at admission and twice a year (April and October). In the assessment, a wage-weighted work unit (WWU) is assigned to the patient. The WWU represents an estimate of the relative quantity of resources used to care for long-term care patients (Schneider, Fries, Foley, Desmond, & Gormley, 1988). Starting in FY98, we used the relative values from the RUG assessments to adjust VA long-term care costs for case-mix.

This section describes the methods using numbers from FY98. The methods are the same for FY99 and FY00, although the numbers are different.

In FY01 - FY04, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are now available, and efforts are underway to improve the quality of these data so that they can be used for research purposes.

6.1 Case mix index

In FY98, there were 45,694 nursing home stays in the VA utilization files. To adjust nursing home costs for case mix, we calculated three case-mix indexes:

(1) Patient level case-mix index
(2) Medical center nursing home case-mix index, which is a mean index for all patients at a medical center weighted by the length of stay
(3) National nursing home case-mix index, which is a weighted mean index of VA nursing home patients

To estimate costs that occurred within FY98, we included only the number of days from October 1, 1997 through September 30, 1998. However, to calculate patient case-mix, we included all possible assessments associated with nursing home stays in FY98 from the following six files:

1. FY972: the admission assessment file in the second half of FY97
2. FY981 the admission assessment file in the first year of FY98
3. FY982 the admission assessment file in the second half of FY98
4. OCT97 the regular assessment file in October 1997
5. APR98 the regular assessment file in April 1998
6. OCT98 the regular assessment file in October 1998

The October 1998 assessment was included because it was the best measure of resource use at the end of the fiscal year.

6.2 Patient level case-mix

The Resource Utilization Group (RUG II) instrument contains 17 levels of resource use in six categories (see table 6.1). In each category, a letter indicates different level of resource use. For example, a patient in Rehabilitation B is assigned 1000 for WWU. In this report, the value of WWU is called RUG score. This assessment information is contained in the Patient Assessment File (PAF) at Austin Automation Center.

Depending on the date of admission and the length of stay, the number of assessments that a patient could obtain during a nursing home bed-section stay varies. In general, every nursing home patient is assessed at admission. We calculated an average RUG score weighted by the number of days between assessments as the case-mix index of a nursing home stay. To calculate an average RUG score for resource use, we were concerned that there was no measure of resource use at discharge. Resource use could change substantially at discharge from the last assessment, especially when the patient died at discharge. Therefore, we developed a regression model to estimate a RUG score at discharge.

Table 6.1: RUG II classification and Wage-Weighted Work Units

RUG Category		WWU
Rehabilitation	A	896
	B	1000
Special Care	A	867
	B	976
Clinically Complex	A	484
	B	711
	C	778
	D	929
Behavioral	A	479
	B	640
	C	744
Physical	A	413
	B	546
	C	640
	D	707
	E	820
CHR VENT DEP		1800

6.2.1 RUG score at discharge

To estimate resource use at discharge, we developed two regression models: a one-assessment and a two-assessment model, depending on the number of assessments per patient. People in the one-point model had one previous assessment, whereas patients included in the two-point model had at least two assessments.

For the models, we selected a sample of nursing home discharges that occurred in either October or April; that is the discharges were within 30 days of the last assessment. Selected cases also had at least three assessments during the stay from all assessment files between fiscal year 1994 and 1999.

We used the last assessment as the dependent variable. The explanatory variables included one or two RUG scores from previous assessment(s), discharge status (died in hospital or alive at discharge), and length of stay. The time between two regular assessments is 180 days. Consequently, we gave an estimated RUG discharge score for those nursing home stays in which the last assessment was more than 90 days (half of the length between two regular assessments) before the discharge. If a patient only had one RUG score, we estimated the discharge assessment based on the one-point model. Otherwise, for nursing home stays with more than two RUG scores, we used the two-point model. In a sensitivity analysis (not shown), we also examined models including more than two RUG scores. The coefficients of RUG scores with more than two lag periods were not statistically significant. The two models (one-point and two-point) are specified below.

Two-point model: patient had at least two assessments and discharge was more than 90 before discharge

$$WWU_d = b_0 + b_1WWU_1 + b_2WWU_2 + b_3D_1 + b_4D_2 + b_5(WWU_1*D_1) + b_6 (WWU_2*D_1) + b_7Died + b_8LOS240 \quad (R^2 = .2940)$$

where

$WWU_{d =}$ Estimated RUG score within 30 days of discharge

WWU_1 and WWU_2 = the last two WWU assessment scores (WWU_1 is the most recent assessment)

D_1 = an indicator (D_1 =1 when $WWU_1 - WWU_2 > 0$)

$D_{2 =}$ an indicator (D_2 =1 when $WWU_1 - WWU_2 < 0$)

Died: an indicator (Died =1 when a patient died at discharge)

LOS240: an indicator for length of stay (LOS240 =1 when the length of stay is less than 240 days)

One point model: patient had only one assessment and the assessment was more than 90 days before discharge.

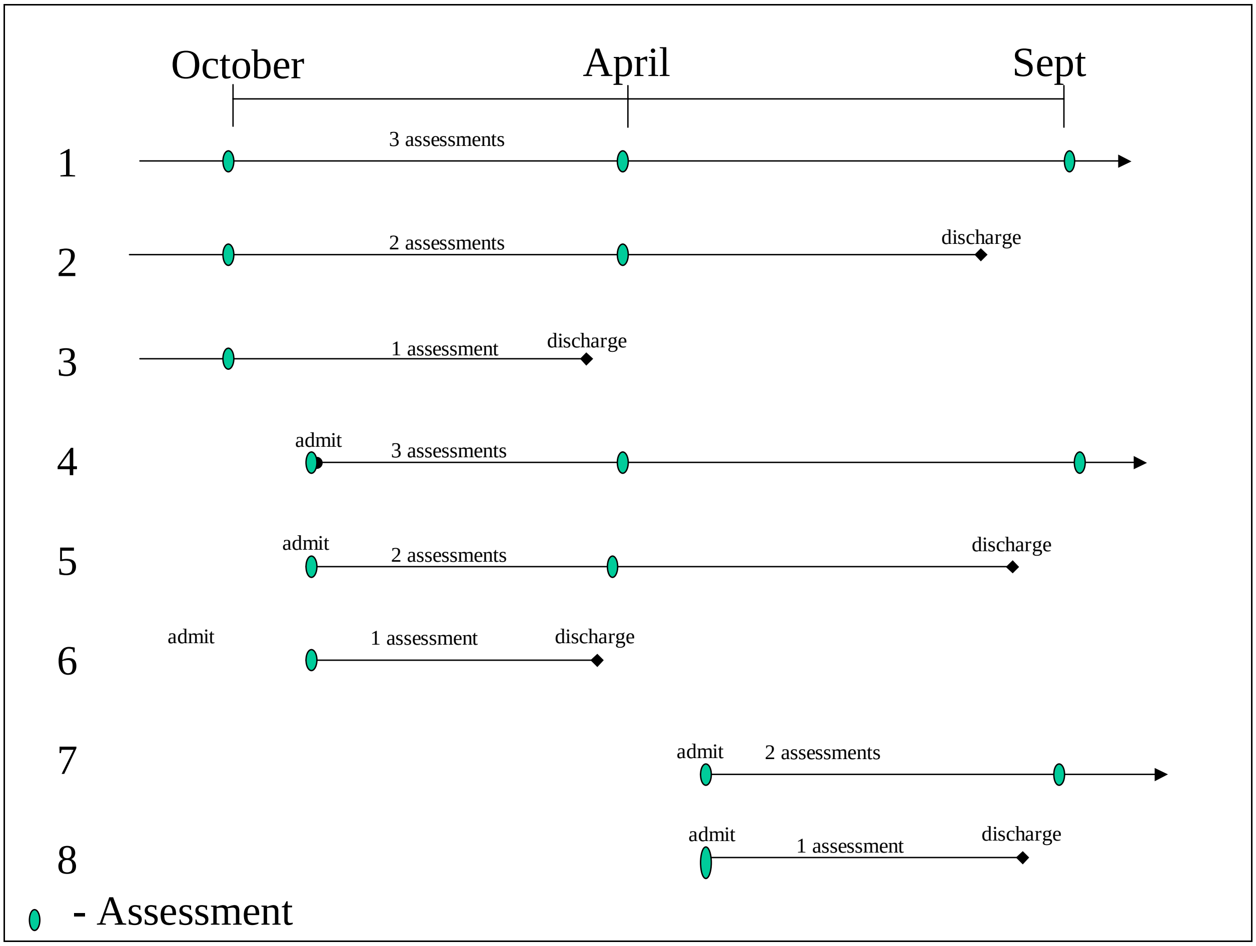
$$WWU_d = b_0 + b_1 WWU_1 + b_2 Died \quad (R^2 = .2411)$$

6.2.2 Average WWU

Based on the admission date and length of stay, patients could obtain different numbers of assessments during a single nursing home stay. Figure 6.1 lists 8 possible combinations of admission and discharge time for a stay within a fiscal year.

An average RUG score (WWU) was calculated based on available assessments using the following formulas for each of the 8 situations in Figure 6.1. It was weighted by the proportion of stay preceded or followed each assessment.

Figure 6.1: Number of possible assessments used to calculate an average WWU in 8 situations



Situation 1

$$\overline{wwu} = \frac{1}{4} wwu_{oct1997} + \frac{1}{2} wwu_{apr1998} + \frac{1}{4} wwu_{oct1998}$$

Situation 2

If the discharge month was between April and June, then

$$\overline{wwu} = \frac{90}{LS} wwu_{oct1997} + \frac{(LS - 90)}{LS} wwu_{apr1998}$$

If the discharge month was between July and September, a RUG score (WWU_d) at discharge was estimated using a regression model and the average WWU score was

$$\overline{wwu} = \frac{90}{LS} wwu_{oct1997} + \frac{(90 + (.5LA)}{LS} wwu_{Apr1998} + \frac{LA}{2LS} wwu_{dis2}$$

LS = the number of days from October 1, 1997 through discharge
LA = the number of days from April 1, 1998 through discharge
WWU_{Dis2} = A RUG score estimated by the two-point model.

Situation 3

If the discharge month was between October and December of 1997, the average RUG score (WWU) was the October 1997 assessment.
If the discharge month was between January and March of 1998,

$$\overline{wwu} = .5(wwu_{oct1997} + wwu_{dis})$$

Situation 4

$$\overline{wwu} = \frac{LA}{2LS} wwu_{adm} + \frac{(90 + .5LA)}{LS} wwu_{apr1998} + \frac{90}{LS} wwu_{oct1998}$$

where
LS = the number of days from admission through September 30, 1998.
LA = the number of days from admission through March 31, 1998.
WWU_{Adm} = RUG score at admission.

Situation 5

If discharge was between April and June of 1998, then

$$\overline{wwu} = \frac{LA_1}{2LS} wwu_{adm} + \frac{(.5LA_1 + LA_2)}{LS} wwu_{apr1998}$$

where
LS = the number of days from admission through discharge,
LA_1 = the number of days from admission through March 31, 1998, and
LA_2 = the number of days from April 1, 1998 through discharge.

If the discharge month was between July and September of 1998, then

$$\overline{wwu} = \frac{LA_1}{2LS} wwu_{adm} + \frac{(LA_1 + LA_2)}{2LS} wwu_{apr1998} + \frac{LA_2}{2LS} wwu_{dis2}$$

where
LS = the number of days from admission through discharge,
LA_1 = the number of days from admission through March 31, 1998, and

LA_2 = the number of days from July 1, 1998 through discharge.
WWU_{Dis2} = A RUG score estimated by the two-point model.

Situation 6

$$\overline{wwu} = .5(wwu_{adm} + wwu_{dis1})$$

where WWU_{Dis1} was estimated by the one-point model.

Situation 7

$$\overline{wwu} = .5(wwu_{adm} + wwu_{oct1998})$$

Situation 8

Same as the formula used for situation 6.

6.2.3 Exceptions

Among the 45,694 nursing home stays in FY98, 891 did not have any assessments from the 6 assessment files we selected. We assigned the medical center average case-mix as the case-mix index for those nursing home stays.

Among the 44,803 nursing home stays with at least one assessment, 1,432 (3%) did not match with assessments within the expected time windows, which starts from 5 days before and 15 days after the admission date. For these nursing home stays, we calculated an average of up to 3 most recent assessments in FY98 as the case-mix index.

6.3 Case-mix index of a medical center

We calculated a case-mix index for each medical center (LWWU) to measure the average case mix of nursing home patients in the medical center. The LWWU is equal to the sum of case-mix adjusted number of nursing home days divided by the total number of nursing home days in the medical center.

$$LWWU = \frac{\sum_{1}^{n} \overline{wwu}_i * LOS_i}{\sum_{1}^{n} LOS_i}$$

where WWU_i is the case-mix index for patient i,
LOS_i is the length of stay for patient i, and
n is the total number of nursing home admissions in the medical center.

6.4 National case-mix index

We also calculated a national average case-mix index (NWWU) by a similar method:

$$NWWU = \frac{\sum_{1}^{n} \overline{wwu_i} * LOS_i}{\sum_{1}^{n} LOS_i}$$

where WWU_i is the case-mix index for patient i,

LOS_i is the length of stay for patient i, and

N is total number of nursing home admissions in all VA nursing homes during the fiscal year.

6.5 Relative Value Unit (RVU)

Case-mix indexes were normalized at the national as well as at the local (medical center) levels. For each nursing home stay, a national RVU (RVUN) was calculated by dividing the individual case-mix index (WWU) by the national average case-mix index (NWWU) and a local RVU (RVUL) was calculated by dividing the individual case-mix index (WWU) by the local average case-mix index (LWWU). The average case-mix adjusted cost was calculated at two levels: the local (medical center) average cost and the national average cost.

<u>6.5.1 Average case-mix adjusted local cost</u>

The average case-mix adjusted local nursing home cost for patient I at the medical center j was calculated by

$LC_{ji} = DC_j \times LOS_{ji} \times RVUL_{ji}$

where:

LC_{ji} - average case-mix adjusted local nursing home cost for patient I at the medical center j,

DC_j - average non-adjusted average per diem cost of the medical center j,

LOS_{ji} - the length of stay for patient I at the medical center j,

$RVUL_{ji}$ - the local RVU for patient I at the medical center j.

<u>6.5.2 Average case-mix-adjusted national cost</u>

The average case-mix-adjusted national cost was calculated as

$NC_i = DC \times LOS_i \times RVUN_i$

where

NC_i - average case-mix adjusted national nursing home cost for patient I,

DC - average non-adjusted national per diem cost,

LOS_i - the length of stay for patient I,

$RVUN_i$ - the national RVU for patient I.

6.6 Distribution of case-mix

The individual RVUs and the medical center average normalized case-mix indexes for FY98 are listed in table 6.2.

Table 6.2: Distribution of RVUs at Patient and Institutional Levels in FY98

Mean	Std	Min	Max
Individuals			
1.0	0.274	0.59	2.74
Medical Center Means			
1.0	0.094	0.82	1.31

Table 6.2 shows that there is a substantial variation in patient case-mix. The maximum RVU is more than 4 times of the minimum. If nursing home costs were not adjusted for case mix, such large differences in resource use would be missed. Also, the average case-mix for medical centers varies considerably. This could be caused by the differences in patients' health status, institutional characteristics, or the quality of assessment measures. Further investigation is needed to understand these patterns.

Chapter summary

- VA long-term care patients are evaluated using the Resource Utilization Group (RUG- II) assessment method.

- These assessments are performed at admission and twice a year (April and October).

- The assessment assigns Wage-Weighted Work Units (WWU) to the patient. The Wage-Weighted Work Unit represents an estimate of the relative quantity of resources used to care for long-term care patients.

- When a patient has more than one assessment, we calculated a weighted average WWU, with weights reflecting the proportion of the stay that proceeded and /or followed each assessment.

- When the most recent assessment was longer than 90 days from the discharge, we estimated a WWU at discharge using a regression model.

- When a nursing home stay did not have any assessment recorded in the Patient Assessment File, we assigned the institutional average RVU to the stay.

- We used the RUG scores (WWUs) to adjust for resource use. This was done by summing together the number of weighted days for patient stays in a medical center. The total cost from the CDR was then divided by the total number of weighted days, yielding a weighted daily cost. To estimate a person's average cost for a stay, we multiplied the daily cost per weighted day by the weight (RUG score) and the length of stay.

- In FY01 - FY04, the cost of long-term care is a per diem rate. In FY01, VA switched from RUG II to the RUG III/MDS dataset. These new RUG scores are available, but more work is needed before they can be used to estimate cost.

Chapter 7. The cost of acute medical-surgical hospitalizations

The cost of acute medical-surgical hospital care in VA can be more accurately estimated by incorporating diagnostic information from the administrative record, and avoid the assumption that every day of stay is of equal cost (Barnett, 1997). We used an econometric cost function, with parameters estimated from non-VA data, to impute the costs for acute medical-surgical stays in the VA.

For more information see, Wagner, T. H., Chen, S., and Barnett, P. G. "Using average cost methods to estimate encounter-level costs for medical-surgical stays in the VA," *Med. Care Res. Rev.* 60 (2003) 15S-36S.

This method relies heavily on non-VA relative value weights. These weights, known as DRG weights, are used to pay hospitals for providing care to Medicare patients. Upon discharge, patients are assigned a Diagnosis Related Groups (DRGs) based on their primary diagnosis. This weighting system is used by the Centers for Medicare and Medicaid Services (formerly the Health Care Financing Administration) to determine Medicare payments to hospitals.

This section presents the cost function that we developed with Medicare data. Given the complexities in this chapter, a flow diagram is provided in Appendix B to help readers visualize the process.

7.1 Making an acute medical-surgical inpatient discharge database

The VA keeps track of bedsections (note: bedsection is a VA-specific term that is most analogous to a hospital ward). Because a patient can get transferred among bedsections multiple times within a single acute medical-surgical hospital stay, keeping track of bedsections provides us with a great amount of detail that is necessary for identifying acute medical-surgical stays.

To use non-VA relative value units, we had to restructure the VA data to use the same definition of acute stays as is found outside the VA. Most non-VA databases are organized as discharge databases with each record representing an acute medical-surgical hospital discharge. While the PTF Main is a discharge database, it does not distinguish between acute medical-surgical and non-medical/surgical care. In addition, the PTF Bedsection file is a discharge file but it separates each record into bedsection stays, even if the bedsections are all part of one acute medical-surgical stay. Therefore, we had to make a database of acute medical-surgical discharges using the PTF bedsection file.

We defined an acute medical-surgical stay based on the following bedsections: 01-12, 14-17, 19, 31, 34, 35, 50-63, 75, 83. Of these, the surgical bedsections are 50-63 and the remainder are acute medicine bedsections. These are the bedsections identified by the VA as the source of workload for costs reported in the acute medical and surgical cost distribution accounts.

We then sorted the data by scrambled social security number (SCRSSN), medical center (STA3N), bedsection in day (BSINDAY) and bedsection out day (BSOUTDAY). Acute medical-surgical bedsection stays that were contiguous in time were considered to be part of the same hospitalization. Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection yielded one non-medical/surgical and two acute medical-surgical stays.

We created a program to accumulate contiguous acute medical-surgical bedsection stays. The program also performs a number of other important functions, such as recalculating length of stay, identifying the highest DRG weight from multiple bedsections (see section 7.2), and calculating number of days spent in intensive care (ICU). The program produces two discharge files, one for acute medical-surgical care and one for non-medical/surgical care. The SAS code for accumulating the stays is available upon request.

7.2 Selecting the DRG and the relative value associated with a DRG

VA assigns a DRG to each bedsection segment of the hospital stay, and another DRG to the PTF main file, representing the DRG for the entire stay. The DRG is based on the principal diagnosis, the condition that is responsible for the patients' admission to the hospital.[5] The Health Care Financing Administration has developed a set of weights based on the DRG (DRG weights). These DRG weights are used to pay hospitals for Medicare patients.

We decided to use the DRG weights for our relative weights in the cost function. DRG weights are not part of the VA databases and were obtained from CMS and added to the VA files. Given that we had 1996 Medicare data, we merged the 1996 DRG weights from CMS with the PTF bedsection file. Then while we were making the acute medical-surgical VA hospital discharge file, the highest DRG weight across all bedsections was maintained. The rationale for this is that a private hospital would follow the same logic to maximize reimbursement.

We considered, but did not use, other relative value systems. We decided that the weights developed by states to pay Medicaid are likely to reflect the patterns of practice in a specific state and that it would not be appropriate to apply them to the VA's national system of hospitals. Some relative value systems, such as the Severity of Illness Index, may provide some additional measure of relative cost (Averill et al., 1992), but they are not feasible for us to implement as they require data that are not available in VA utilization data at Austin. Patient Management Categories and Disease Staging are case-mix methods that can be applied to standard datasets, but they have been found to explain only 1-2% more variation than DRGs used alone (Calore & Iezzoni, 1987).

---

[5] Prior to October 1, 1994, VA used the primary diagnosis to define DRGs. The primary diagnosis is the most important condition treated in the stay (as opposed to the principal diagnosis, which is the diagnosis responsible for the patient's admission to the hospital). VA DRGs from stays that ended prior to this date are thus not strictly comparable to non-VA DRGs from that time period, which have always used principal diagnosis.

We use the appropriate DRG weight file from the Centers for Medicare and Medicaid Services (CMS) for each year of the Average Cost data. For the FY04 cost estimates, we used the 2004 DRG weight file from CMS.

7.3 Length of stay

Length of stay is reported in the PTF bedsection file. But we had to recalculate length of stay according to our definition of acute medical-surgical stay (see section 7.1). Consequently, length of stay represents all days the patient spent in contiguous acute medical-surgical care bedsections during the stay.

7.4 Building the cost function

In past years we used an econometric method of estimating VA acute medical-surgical care costs (Barnett, 1997). Starting with FY98, we developed a cost function for estimating the cost of acute medical-surgical care. The cost-function is based on non-VA data, where the hospital stay as the unit of analysis. Using the stay (rather than the average stay) as the unit of analysis provides much more variation, including observations with high DRG weights and long lengths of stay. The cost function approach allowed us to construct a more complex model that better simulates the cost of stays with characteristics that are very different from the mean.

While the mechanics of the cost function are complicated, the intuition is relatively straightforward. We built a statistical model with a hospital discharge dataset. This regression model had cost adjusted charges on the left-hand side. On the right-hand side, we included variables such as length of stay, DRG weight, whether the patient died in the hospital, age, gender, and so forth. We saved the parameters from the regression model (i.e., the beta coefficients). This vector of coefficients was used to estimate costs in the VA data. It is important to note that the only way this approach can work is for both datasets to have the exact same right-hand side variables.

7.4.1 Data

We chose to use Medicare data for the cost function. Medicare data have some limitations, namely that Medicare does do not cover non-disabled individuals under age 65. For this reason, we carefully compared Medicare data from veterans to the Health Care Cost and Utilization Project (HCUP) data.

To provide some background on these datasets, the Medicare data were a subset of the 1996 MedPar file. The MedPar file was constructed by researchers at the Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC). They established a cohort of all veterans who were users of either inpatient or outpatients VA services between 1992 and 1994 and who had their 65th birthday in 1994. This cohort was then linked to Medicare denominator file to obtain Medicare enrollment. The file that we received represented 372,046 stays from hospitals in the continental US.

The HCUP data represents discharges from all types of hospitals in 22 states. Detailed information on the HCUP dataset is available on-line from www.ahrq.gov.

The primary question is, can we use the Medicare data to build a model that can estimate costs for younger veterans? Recall that Medicare data do not include non-disabled individuals under age 65. We answered this question by building a cost function with Medicare data. The function was then used to estimate the cost of stays in the HCUP sample. We then compared the estimated Medicare costs to the costs reported in the HCUP. This comparison was made for adults over 65 as well as adults under age 65. The remainder of this section describes this comparison.

First we selected a 40% random sample of non-ESRD Medicare claims in the MAVERIC cohort (125,457). With these claims, we estimated the following model:

CAC=a+b1died +b2sex +b3age+ b4npr+ b5npr2 +b6los + b7poslos + b8neglos + b9nlos2 +b10plos2 + b11nlos3 + b12drgwt +b13drgwt2 +e

where

CAC is cost adjusted charges
npr is number of surgical procedures
npr2 is number of surgical procedures squared
los is DRG specific length of stay
poslos is (average los-los) if average los > los
neglos is (average los-los) if average los < los
nlos2 nlos3 are square and cubic terms of neglos
plos2 is squared term of poslos
drgwt is CMS drgwt
drgwt2 is drgwt squared

The parameters from this model were saved and then used them to impute estimated costs for HCUP. We tried alternative model specifications, including the log transform of cost adjusted charges and excluding people with end stage renal disease (ESRD). In all of these alternative specifications, the parameters for the older people were remarkably similar to the parameters for the younger populations. We concluded that we could use the Medicare data to estimate the costs of younger hospitalized patients. The main advantage to this approach is that the Medicare data identify the number of days spent in intensive care (ICU). Because intensive care units are resource intensive and costly, being able to estimate this parameter was a key advantage.

For the FY01 - FY04 cost estimates, we used the 1999 MedPar file of veterans for estimating costs.

7.4.2 Cost adjusted charges

Utilization databases, like the Healthcare Cost and Utilization Project (HCUP) or Medicare, report charges incurred in a hospital. Yet, it is generally known that health charges usually exceed the cost of providing care. However, the degree to which charges exceed costs is not completely random. Hospitals and medical centers are somewhat idiosyncratic in how they generate bills.

Hence, we want to adjust the charges for two reasons: (1) to deflate charges so that they more closely reflect costs, and (2) to remove hospital specific idiosyncrasies. The ratio of costs to charges (RCC), described in detail below, is one way of making this adjustment.

Adjusting charges with the RCC leverages information that every hospital annually reports to Medicare in the Medicare Cost Report. The Medicare Cost Report is a very large report that hospitals are required to complete if they want to receive federal reimbursement.

In the Medicare Cost Report, there are variables for each hospital's total charges and total costs. In the most recent Medicare Cost Report (PPS version 13), the field for charges is 2135 and the field for costs is 2138. We extracted these fields along with the hospital's Medicare identification number (PPS number). The quotient (i.e., the result of dividing costs by charges) was the ratio of costs to charges (RCC). The RCC usually ranges between 0.5 and 1.0. To actually adjust charges, the RCCs were linked to the Medicare dataset with the PPS number. The charge data were then adjusted by the RCC.

For example, if we want to use the RCC to adjust charges in a dataset, such as the HCUP dataset, we must first crosswalk the RCC dataset to the HCUP dataset. This can be a complicated process, especially for crosswalking the HCUP to Medicare (for details, see http://www.herc.research.med.va.gov/resources/faq.asp). Once we crosswalk the files, we then multiply charges by the RCC. Recall that the RCC is a hospital-specific adjustment. In other words, within any given hospital the RCC will be constant.

For FY01 - FY04, we obtained the 1999 Medicare Cost Report (PPS16). PPS16 has different variables than the PPS13. They provide department level costs and charges. We used this to create a facility cost to charge ratio.

7.4.3 The dependent variable

We used cost adjusted charges as our dependent variable when we built the cost function. However, the cost adjusted charges from the Medicare data are not normally distributed. Because of the skewness, we tried transforming the cost adjusted charges. While the log transformation helped reduce the appearance of skewness, the non-logged function consistently performed better than models with logged cost adjusted charges. Using logs presents additional hurdles because the estimated costs need to be transformed back to the original metric (dollars), adjusting for retransformation bias. The usual adjustment for retransformation bias is the smearing estimator (Duan, Manning, Morris, & Newhouse, 1983). While relatively simple to implement, this adds another layer of complexity to the entire process.

7.4.4 Length of stay

There are different ways to include length of stay in a cost function. The most obvious way is to include it without making any transformations, such that length of stay is a positive integer. Variations on this approach were also considered, such as a set of dummy variables representing different lengths of stay.

A second method for including length of stay involves comparing the patient's length to the average length of stay for all patients with that DRG. This second approach requires knowing the average length of stay for each DRG. This information is conveniently provided by CMS with the DRG weight file. We found slight advantages to the second approach as the transformation turned the length of stay from a positive integer into a continuous scale. Having a continuous scale provides slightly more ability to discriminate costs based on deviations in length of stay.

We used the second approach. In addition, we relaxed the constraints of our earlier estimates, allowing the cost of marginal days of stay to vary, depending on the length of stay.

Note that we examined only those records of patients discharged during the fiscal year under study. We included days of stay in acute medical-surgical bedsections, even if they occurred in previous fiscal years, and excluded data from stays that were not complete by the end of the fiscal year. This is distinct from the rest of our method, which considered only the days of stay that occurred during the fiscal year under study. We also calculated the length of stay in ICU bedsections. For each acute medical-surgical hospital stay, we found the number of days spent in the medical and surgical ICU bedsections.

7.4.5 Individual DRG intercepts or DRG weights

We found little marginal value in including dummy variables for each DRG. When we included DRG weight (squared and cubic terms), the gain in R^2 was less than 1%. Given the additional complexity in estimating this model, we decide to not use it. Instead, we decided to use DRG weight in our cost function along with the DRG weight squared and cubed. In the final model, we also interacted the Medicine Major Diagnostic Category (MDC) and Surgery MDC with length of stay.

7.4.6 Final model

The final cost function model based on a 50% sample of the Medicare data is shown in Table 7.1. The variable definitions follow.

7.4.7 Outliers

Outliers can have undue leverage on a regression model. After we ran the model, we found that the model fit the data reasonably well. However, the fit was based primarily on the high cost users. The model did not fit as well for low-cost users, due in part to heteroskedasticity.

One solution involves removing or “trimming” outliers. We tried this and retested the model fit. Our methods and findings are below. We first identified outliers by using the Medicare outlier designation (n=1880). This did not help the fit of the model with low-cost cases because the outlier designation typically identifies the expensive cases.

Then we empirically identified outliers by generating Cooks' distance. Cooks' distance is the leverage of case *I* on the OLS regression coefficients (βhat). It can be thought of as an F test comparing the beta coefficients with and without observation *I (i.e.,* βhat to βhat_{-I}). Large values for Cook's distance suggest that the case has a lot of leverage.

We trimmed outliers in our regression models using three exclusion criteria:[6]

1) Cooks distance >0.001 (excluded 968 observations, ~0.8%)
2) Cooks distance >0.0001 (excluded 2,101 observations, ~1.7%)
3) Cooks distance >0.00001 (excluded 8,431 observations, ~6.6%)

We found that we could estimate better fitting models if some outliers were excluded. This gain was mainly within the lowest quartile of costs. Table 7.2 presents correlation coefficients between actual cost adjusted charges (CAC) and estimated cost adjusted charges. Note, however, that not always did removing more outliers lead to a better fitting model. In quartile 1, only model #3 yielded higher correlations.

[6] We also compared logged CAC models. In every case, the log models fit significantly worse and yielded much larger differences between estimated costs and actual costs.

Table 7.1: Full model based on 50% random sample of Medicare data (FY98-00)

Source	SS	df	MS
Model	3.8009e+13	27	1.4078e+12
Residual	1.3554e+13	321555	42152405.8
Total	5.1564e+13	321582	160343662

Number of obs	321583
F(27,321555)	33396.73
Prob > F	0.0000
R-squared	0.7371
Adj R-squared	0.7371
Root MSE	6492.5

cac	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
died	2671.211	57.21167	46.690	0.000	2559.077	2783.344
sex	32.90875	61.21531	0.538	0.591	-87.0715	152.889
age	-34.22324	1.851834	-18.481	0.000	-37.85278	-30.5937
ndx	619.0444	81.09738	7.633	0.000	460.0959	777.993
ndx2	-146.7017	16.61743	-8.828	0.000	-179.2714	-114.1321
ndx3	10.97541	1.022981	10.729	0.000	8.970401	12.98043
los	104.255	9.083375	11.478	0.000	86.45187	122.0582
poslos	670.9503	10.10664	66.387	0.000	651.1415	690.759
neglos	182.4991	29.68224	6.148	0.000	124.3228	240.6755
nlos2	-109.8903	7.980714	-13.769	0.000	-125.5323	-94.24832
plos2	-.7170458	.021736	-32.989	0.000	-.7596478	-.6744437
nlos3	-4.587643	.5484962	-8.364	0.000	-5.66268	-3.512606
plos3	3.32e-06	.0000198	0.168	0.867	-.0000354	.000042
drgwt	4860.036	63.69243	76.305	0.000	4735.201	4984.871
drgwt2	-255.1638	11.0401	-23.112	0.000	-276.8021	-233.5255
drgwt3	12.97284	.5057919	25.649	0.000	11.98151	13.96418
surg	1069.883	78.21631	13.679	0.000	916.581	1223.184
surlos	-42.31538	11.16155	-3.791	0.000	-64.19169	-20.43906
pl_sur	421.5315	15.61753	26.991	0.000	390.9216	452.1415
nl_sur	328.304	36.252	9.056	0.000	257.2511	399.3569
pl_sur2	-1.384451	.1793446	-7.720	0.000	-1.735961	-1.03294
pl_sur3	.001167	.0006719	1.737	0.082	-.00015	.002484
nl_sur2	47.49814	8.419396	5.642	0.000	30.99636	63.99991
nl_sur3	3.636805	.55208	6.587	0.000	2.554745	4.718866
icudays	593.0367	7.165874	82.758	0.000	578.9918	607.0816
icudays2	10.27421	.2713893	37.858	0.000	9.742298	10.80613
icudays3	-.0325464	.0017843	-18.240	0.000	-.0360436	-.0290492
_cons	413.7664	181.3739	2.281	0.023	58.27884	769.254

We decided not to remove outliers because we realized any decision about which outliers should be removed would be arbitrary and would affect the model's fit. The full model fits almost as well (and better in some instances), therefore we saw little rationale for removing outliers. Table 7.2 also shows how well the model predicts costs with the other 50% of the data (out of sample). In many cases, the out-of-sample predicted costs are quite close to the actual Medicare costs. As is shown in Table 7.1, the overall R^2 of the model is approximately 0.74.

Table 7.2: Correlations between estimated costs and actual costs for the full model and for three outlier restricted models

	Actual costs							
	Quartile 1: <$2605		Quartile 2: $2605<cac<$4484		Quartile 3: $4484<cac<$8472		Quartile 4: >$8472	
	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample	In sample	Out of sample
Sample size	*38304*	*38144*	*39167*	*38594*	*39939*	*40801*	*43348*	*43286*
Model with all cases	correlation coefficients							
estimated costs	0.126	0.190	0.301	0.291	0.389	0.357	0.814	0.808
Restricted models								
(1)	0.057	0.204	0.309	0.005	0.396	0.250	0.641	0.699
(2)	0.071	0.209	0.313	0.011	0.398	0.279	0.718	0.749
(3)	0.185	0.202	0.313	0.305	0.393	0.392	0.769	0.775
Model estimated with log(CAC)	0.083	0.109	0.303	0.290	0.390	0.381	0.389	0.106

Notes: (1) cost function was estimated excluding cases with a cooks' distance >.001 (least restrictive)
(2) cost function was estimated excluding cases with a cooks' distance >.0001 (more restrictive)
(3) cost function was estimated excluding cases with a cooks' distance >.00001 (most restrictive)

7.5 Observation days

Beginning in 1997, VA created 7 new codes for observation bedsections to report inpatient care provided in observation units. Most stays involving these codes are recorded in the observation PTF files, which is a new set of files in the PTF. These stays, even if there are associated with an inpatient record in the Acute PTF file, are kept in a separate observation bed file at Austin. The structure of the observation files mirror the PTF inpatient files. We found that many stays reported in this file precede or follow stays in the acute medical-surgical PTF file. When calculating length of stay, some analysts will want to regard these observation days as part of acute medical-surgical stays.[7]

[7] Nearly 73,000 days of stay were assigned to observation bed sections in FY99 (out of 13.5 million days in VA hospitals). Most of the observation stays were one day long, but this was not always the case. Most observation days were in medicine, surgery, and psychiatry observation bedsections. We recently examined the FY99 data and found that 19,428 (26%) of the observation stays immediately preceded a stay reported the PTF bedsection files. Another 319 observations stays followed stays in the bedsection file. (Our analysis was limited to PTF bedsection file. It is also possible that observation stays precede or follow stays reported in the PTF extended care file.)

For the cost of observation bed stays, for FY98 onward we costed each day at the marginal cost of an additional day (i.e., $684; see section 7.6). This method may underestimate the cost of stand-alone observation stays. Alternatively, it may overestimate the cost of an observation stay that preceded a hospitalization. We hope to develop and test new methods for costing observation bed stays in the future.

7.6 Negative or implausible costs

After estimating FY98 VA costs with the cost function (see Table 7.1), we found that the function had imputed negative costs for 2,974 of the 541,567 (0.6%) acute medical-surgical hospitalizations. This is because the cost function was not constrained to predict non-negative estimates. Therefore, rare combinations of right-hand-side variables can lead to negative predictions. These 2,974 records were assigned the cost of a marginal day of stay ($684.75).

The cost of a marginal day of stay was calculated in a simulation with the 1996 Medicare data. Adjusting for all other covariates in a linear regression, we identified the cost for an additional day of stay. Holding all other factors at their mean, if a person stayed an additional day, they had an additional $684.75 of cost adjusted charges.

While some stays were not assigned negative costs, they were given very low costs. For instance 42 hospital stays had positive costs less than $5. We decided that any stay with a cost less than $684.75 was implausibly low and an artifact of the cost function. By setting this rule, it effectively set a floor on the estimated cost per stay. A total of 9,632 (2%) cases had non-negative costs less than $684.75. These cases were all given $684.75 per day (86% had a length of stay of one day). In the future, we will explore other methods for determining the cost of these cases, including setting constraints on the cost function.

7.7 Reconciling to the CDR

The cost function is based on non-VA relative value weights and non-VA cost adjusted charges. The estimated costs must be reconciled to the Cost Distribution Report to reflect VA costs. Reconciliation can happen at many levels including the department, medical center, and nationwide. We chose to reconcile the estimated costs to the medical center and nationwide; we decided not to reconcile the estimated costs to the department. Given that the CDR and PTF are not reconciled against each other, our concern was that there would be too much variability in department-level costing.

Reconciling the costs to the medical center results in “local” cost estimates, while reconciling the costs for the entire VA results in “national” cost estimates. Therefore, this process results in the creation of 2 VA cost estimates: a local cost estimate (costl) and a national cost estimate (costn).

The logic behind reconciling the costs is straightforward. For the local cost estimate we sum together the estimated costs for a medical center and divide this amount by the total acute medical-surgical care CDR costs (acute medicine and surgery) for the medical center. The quotient of this division is a scaling factor. By multiplying the estimated cost by this scaling factor, we ensure that the sum of the estimated costs is equivalent to the CDR costs.

Unfortunately, the reconciliation is easier said than done. Recall that the CDR reports costs for the fiscal year while the acute medical-surgical hospitalization data represent discharges. For FY98 data, some stays that ended in FY98 started before FY98. At the same time, there were people hospitalized in FY98 who were still in the hospital at the end of the fiscal year and are not reported in the FY98 PTF data. To illustrate this point, Figure 7.3 shows the hospitalization that cross the fiscal years. Cases B, C, and E all cross the fiscal years. It is not correct to assume that the cases crossing from FY97 to FY98 are equivalent in number to those cases crossing from FY98 to FY99. Due to the declining trend in inpatient hospitalization, C and E are more common than B.

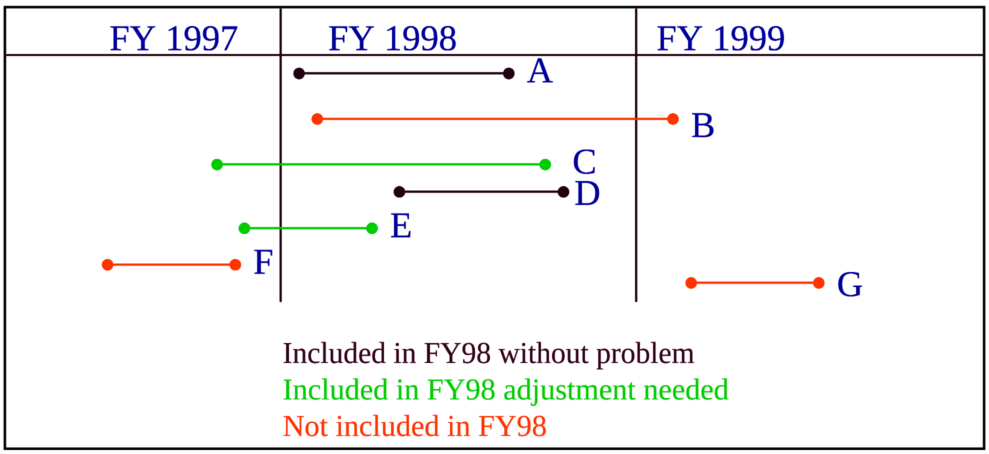


Figure 7.3: Difference between FY view and discharge view Note: A & D are in the med/surg file and need no adjustment
C & E are in the med/surg file and need adjustment
B, G, and F are not in the med/surg file

If no adjustment were made for this fact, then we would overestimate the number of hospitalizations, and thereby underestimate the cost of care per hospitalization. Our correction for this was to adjust the cases discharged in the fiscal year that started before the fiscal year. The FY98 adjustment factor was found by comparing the FY98 Census to the FY 97 Census (see Table 7.3).

After adjusting the discharge data so that it better represented the FY costs in the CDR, we reconciled the estimated costs. The national scaling factors are listed in Table 7.3. We multiplied every estimated cost by this scaling factor to obtain the national VA cost. This ensures that if every acute medical-surgical hospitalization discharged in the fiscal year were summed together that the total would equal the CDR costs.

Table 7.3: Fiscal year adjustment and scaling factors

Fiscal year	Fiscal year adjustment	National scaling factor
FY98	0.93	1.27
FY99	0.9821	1.29
FY00	0.9290	1.41
FY01	1.0442	1.21
FY02	0.9117	1.20
FY03	1.029	1.21
FY04	0.999	1.26
FY05	0.9059	0.567

7.8 Stability of the cost function over time

The cost function for FY98-FY00 was built using 1996 Medicare data. For FY01 - FY04, we used 1999 Medicare data. For FY05, we used 2003 Medicare data. One question is whether the cost-function is robust to the input data that are being used. To answer this question, we used 1994 and 1995 MedPar data that was similar to the 1996 MedPar data. We then ran the identical cost function on all three datasets. The model coefficients from the three datasets were compared. Finally, using the regression model for each year of data, we predicted costs in 1996, using the MedPar 1996 as the criterion. We compared the estimated costs to see if differences would have occurred had they been estimated with 1994 or 1995 MedPar data.

The regression coefficients for all three models were extremely similar (Table 7.5). The predicted costs from the three models were also highly correlated (>0.99; Table 7.5). The results suggest that the cost function is highly robust to the year from which the MedPar data are used.

Table 7.4: Stability of regression coefficients with 1994, 1995 and 1996 MedPar data

	1994		1995		1996	
	coeff	t-stat	coeff	t-stat	coeff	t-stat
died	2837.70	42.410	2803.32	42.650	2671.21	46.690
sex	-41.01	-0.560	-28.73	-0.400	32.91	0.540
age	-42.29	-18.590	-44.42	-19.720	-34.22	-18.480
ndx	250.36	2.740	433.47	4.710	619.04	7.630
ndx2	-80.63	-4.190	-117.71	-6.180	-146.70	-8.830
ndx3	7.44	6.150	9.60	8.120	10.98	10.730
los	50.63	4.660	52.01	4.890	104.26	11.480
poslos	656.08	54.620	666.76	54.250	670.95	66.390
neglos	272.94	9.400	338.59	11.140	182.50	6.150
nlos2	-72.45	-11.940	-71.91	-10.220	-109.89	-13.770
plos2	-1.31	-54.080	-0.62	-10.450	-0.72	-32.990
nlos3	-1.41	-4.830	-1.85	-4.490	-4.59	-8.360
plos3	0.00	30.680	0.00	2.900	0.00	0.170
drgwt	4477.58	58.500	5149.17	69.610	4860.04	76.300
drgwt2	-161.85	-12.100	-325.22	-25.390	-255.16	-23.110
drgwt3	8.02	13.030	16.71	28.480	12.97	25.650
surg	470.37	5.280	526.47	5.890	1069.88	13.680
surlos	-48.96	-3.770	-23.43	-1.810	-42.32	-3.790
pl_sur	416.50	26.280	379.25	22.240	421.53	26.990
nl_sur	222.54	5.670	152.01	3.850	328.30	9.060
pl_sur2	-1.21	-24.520	-0.95	-8.300	-1.38	-7.720
pl_sur3	0.00	18.310	0.00	-1.250	0.00	1.740
nl_sur2	18.26	2.590	3.07	0.390	47.50	5.640
nl_sur3	0.58	1.900	0.72	1.710	3.64	6.590
icudays	395.04	47.070	553.12	67.840	593.04	82.760
icudays2	18.93	58.260	9.29	31.130	10.27	37.860
icudays3	-0.08	-37.720	-0.02	-11.440	-0.03	-18.240
_cons	1819.08	8.640	1416.06	6.650	413.77	2.280

Table 7.5: Pair wise Correlations in predicted costs compared to 1996 costs adjusted charges

	cost94	cost95	cost96
cost94	1		
cost95	0.993	1	
cost96	0.997	0.996	1
CAC 1996	0.856	0.855	0.859

Note: CAC is cost adjusted charges

Chapter summary

- To estimate the costs of acute medical-surgical care for FY98-FY05, we developed a cost-function from Medicare MedPar data restricted to veteran users.

- HERC developed a VA acute medical-surgical dataset using the PTF bedsection file. Contiguous acute medical-surgical bedsection stays were aggregated into a single record. This program also recalculates LOS, ICU days, and keeps the highest DRG weight for all acute medical-surgical bedsection stays.

- In building the cost function, we compared the HCUP dataset to a veteran-restricted Medicare dataset. The Medicare dataset was able to predict the costs of younger people in the HCUP dataset and it identifies ICU days, which are a useful indicator of resource use. Therefore, we used the veteran-restricted Medicare dataset.

- Medicare reports charges. We adjusted the reported charges with a hospital-specific ratio of costs to charges. This deflates the reported charges and removes some hospital-specific billing differences.

- Length of stay was entered into the model as the deviation from Medicare's expected length of stay for that DRG.

- After comparing alternative models, we decided to use DRG weight as the measure of relative weight, rather than allow each DRG to have its own intercept.

- The 1996 MedPar model had an R^2 of 0.7371. The 1999 MedPar model was 0.7539.

- We explored whether to trim influential outliers. This affected the model's fit, and not always positively. Because the cut-off for selecting the outliers was arbitrary, we included all cases.

- For each observation day, we estimated its cost at the marginal cost per day, which we estimated at $684.75.

- The cost function yielded some negative and implausible costs. We set $684.75 (the marginal cost of a day), as the minimum cost possible.

- We reconciled the estimated costs to the CDR for the medical center and the nation. This yielded a local cost estimate (costl) and a national cost estimate (costn).

Chapter 8. User's Guide

This chapter discusses how to use HERC's average cost dataset. The chapter is broken into four sections: 1) a brief summary of the methods, 2) assumptions underlying the dataset, 3) how to correctly use the dataset, and 4) when not to use the dataset. We strongly feel that every user of these data should be knowledgeable in these areas.

Although we hope that these data will be useful, we do not expect that these data will be appropriate for every study. For this reason, later in this chapter we discuss limitations with these data and instances for which these data are not appropriate.

8.1 Summary of methods

8.1.1 Categories of inpatient care

Starting in FY98, we categorized inpatient care into eleven categories: 0) acute medicine, 1) rehabilitation, 2) blind rehabilitation, 3) spinal cord injury rehabilitation, 4) surgery, 5) psychiatry, 6) substance abuse care, 7) intermediate medicine, 8) domiciliary, 9) nursing home care, and 10) psychosocial residential rehabilitation programs (PRRTP). These categories are defined by bedsection (see Table 4.3). While PRRTP care is defined by bedsection, it is only available at approved medical centers. If a non-approved medical center had dollars or days in PRRTP bedsections, these were allocated back into psychiatry and substance abuse care, respectively.

8.1.2 Acute medical-surgical care

Of the eleven categories of care, acute medicine and surgery comprise the acute medical-surgical care. For patients receiving this type of care, we estimated costs using a cost-function from Medicare MedPar data restricted to Veteran users (see Chapter 7). To do this, we developed a VA acute medical-surgical dataset using the PTF bedsection file. Contiguous acute medical-surgical bedsection stays were aggregated into a single record.

In building the cost function, we used a veteran-restricted Medicare (MedPar) dataset. We adjusted the reported Medicare charges with a hospital-specific ratio of costs to charges. In the cost function, length of stay was entered into the model as the deviation from the expected length of stay for that DRG. We also used DRG weight as the measure of relative weight, rather than allow each DRG to have its own intercept.

For each observation day in an acute medicine or surgical bedsection, we costed it at the marginal cost per day, which we estimated at $684.75. The cost function yielded some negative and implausible costs. We set $684.75 (the marginal cost of a day), as the minimum cost possible.

Lastly, we reconciled the estimated costs to the CDR for the medical center and the nation. This yielded a local cost estimate (costl) and a national cost estimate (costn).

8.1.3 Nursing home care

For FY98-FY00, nursing home costs reflect case-mix. Using the Resource Utilization Groups (RUGs) that are collected biannually on nursing home patients, we imputed the daily cost per RUG unit. To obtain the patient's cost per stay, we multiplied each patient's rug score by the per rug cost times the length of stay. The methods for this are discussed in detail in Chapter 6. Nursing home costs for FY01-FY04 were based on an unadjusted per diem. In FY01, VA started using the RUG/MDS data collection tool, rather than the RUG II score. The RUG III data are not yet available.

8.1.4 Non medical/surgical categories

All remaining cost categories were estimated as a daily rate. The total CDR costs were divided by the total units provided in the PTF bedsection file. The daily rate methods are described in detail in Chapter 5.

8.2 Assumptions in the average cost dataset

Throughout this document we have tried to identify assumptions underlying the creation of the acute medical-surgical and non medical/surgical datasets. Both datasets reconcile to the CDR at the level of the medical center and the nation. Costs excluded from the CDR are also not included in our estimates. These include, importantly, the cost of financing capital expenditures and malpractice costs. Our average cost estimates do include indirect costs and physician costs. Table 8.1 shows the included and excluded costs.

Table 8.1: Included and excluded costs

Type	Notes
Excluded	
Capital financing costs	Not included, but this may be noteworthy (5%).
Malpractice expenses	Not included.
Contract provider costs	Excluded are contract services because these costs are not accurately associated with units of care
Community nursing home costs	We excluded cases that were in bedsection 80 with Statyp 42.
Headquarters costs	Excluded are the costs associated with VA headquarters
Prosthetics	Inpatient prosthetics billed separately are not included in the CDR accounts
Included	
Costs for physician services	These costs are included in the CDR. For every stay, physician costs are proportionate to the hospital costs.

Research & education	Included to the extent supported by the VA medical care appropriation.
Indirect costs	We assigned indirect costs to each CDA in proportion to its share of the total direct costs of its group of CDAs.

8.2.1 Data used in the cost function

The average cost estimates for acute medical-surgical stays were based on a cost function that was constructed with Medicare data. The cost function for FY98-FY00 was built using 1996 Medicare data. For FY01 - FY04, we used 1999 Medicare data. For FY05, we used 2003 Medicare data. The Medicare data represented veteran users; excluded were cases in Hawaii, Alaska and cases related to labor and delivery. In using the Medicare data we assumed that the underlying accounting systems for non-VA hospitals could be used to impute estimates for the VA. These imputed estimates were then reconciled with the CDR. If you were to sum all of our cost estimates for a medical center in a given year, you will find that the local cost total is equivalent to the amount posted in the CDR.

8.2.2 The cost of observation stays

Observation stays are a relatively new type of service provided in the VA. There is no analogous type of service provided in the private sector. To estimate the cost of the observation bed stay, we estimated a marginal daily rate and multiplied this times the length of stay. Most people stay in the observation bed for one day; a few outliers stay longer and in these cases, the cost is equivalent to this rate times the length of stay. To calculate the daily rate for observation bed stays, we developed a regression model using Medicare data. With the regression model, we simulated the marginal cost at the mean of data. We then predicted the cost if the person stayed one day longer than the mean. The difference between these two estimates was $684.75. We used this as the daily rate for the observation bed stays.

8.2.3 Costs for high and low-cost procedures

The cost function used to estimate acute medical-surgical costs was presented in chapter 7. As was mentioned in that section, the model does a better job estimating high cost stays. The accuracy of the average cost estimate is better with high-cost cases than with low-cost cases. If you are assessing cases that typically have very low costs, then the average cost provided in the HERC dataset may be inappropriate.

8.2.4 Implicit trimming of outliers

A byproduct of using the cost function is that it removes outliers. Recall that the cost function is a linear regression model. When we calculated the cost for the VA we used the regression model to estimate costs based on averages. If you are interested in high or low-cost outliers, then the HERC dataset may be inappropriate for your use.

8.2.5 Model estimates and negative costs

Another byproduct of using a cost function is that after we imputed the VA costs we had some cases with negative or implausibly low costs. Clearly, a stay cannot have a negative cost. Therefore, we decided that we would set a floor. Any choice of a floor is somewhat arbitrary, but we chose the floor to be $684.75. Recall that $684.75 is the average cost of an additional day of stay (see chapter 7). A total of 12,731 cases had an estimated cost of less than $684.75. For all these cases, we assigned them a cost of $684.75. This cost was their total cost, NOT a daily rate. Of these cases, 83.5% (10,636) had only one day of stay. Another 14% and 2% had a stay of two and three days, respectively. The remainder (101 cases) had up to 8 days of stay; however, there were three outliers who had more than 1000 days of stay. Clearly a cost of $684.75 is inappropriate for someone who stayed 1150 days in the hospital, but we did not make adjustments for these three cases. When you use these costs, compare the length of stay to the cost. Make sure that these three cases are not in your data. If they are, you probably want to exclude them or assign them a different cost.

8.2.6 VISN administrative costs

Each of the VISNs incurs administrative operating costs. We have included these costs under the assumption that they cover coordination expenses required for a large health provider. In the CDR, these costs are assigned to a single medical center within the VISN. From our perspective, these costs should be distributed to all medical centers in the VISN. We are looking into ways of distributing these costs, but for FY98-FY03, these costs remain where they were assigned. This may partly explain deviations in the local costs. This provides a reason for using national costs, but if your study requires local costs, then use them carefully.

8.3 Using the average cost dataset

At Austin, we have provided three datasets. These datasets are listed in Table 8.2 and described below. All of the files can be found in the RMTPRD.HERC.SAS directory.

Table 8.2 The three average cost datasets for FY98

Dataset	Includes	Excludes
dischgXX	• All persons admitted since FY98 and discharged in fiscal year. • Costs for acute medical-surgical are combined with non medical-surgical costs when bedsection stays within a discharge are contiguous.	• stays not completed by end of fiscal year • stays admitted before beginning of FY98 (10/1/97)

mdsrgXX	• All persons discharged from an acute medical-surgical bedsection in fiscal year	• Non medical-surgical bedsections • People who were still in the hospital at end of FY.
nmdsrgXX	• The cost of care provided in rehabilitation, mental health or long-term bedsections during the fiscal year.	• The costs of care provided before the fiscal year are excluded.

Table 8.3 Using the three average cost datasets

Dataset	Sort and merge using	Merge data to
dischgXX	SCRSSN, ADMITDAY, DISDAY, and STA3N.	PTF main files (PM, XM and PMO)
mdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO); BUT must first aggregate the bedsection file
nmdsrgXX	SCRSSN, ADMITDAY, DISDAY, STA3N, BSINDAY, and BSOUTDAY.	PTF bedsection files (PB, XB, PBO), and PTF census files.

8.3.1 Discharge dataset

The discharge dataset was generated by combining the acute and nacute datasets. It represents a discharge dataset, such that it only has cases that were discharged. In addition, only people admitted since the beginning of FY98 are included in the discharge datasets. Patients that were admitted prior to FY98 are excluded

The discharge dataset includes additional variables that track cost subtotals, length of stay subtotals, DRG weight, and ICU days.

Discharge dataset

scrssn	Numeric field. Identifies a patient's scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center's station number. These can change when facilities merge.
adtime	Admission time for an inpatient stay.
admitday	Admission day for an inpatient stay (SAS date)
disday	Discharge day for an inpatient stay (SAS date).
b4fy98	Flag that identifies inpatient stays that began prior to FY98. The numeric variable is either 0 or 1. No HERC costs for these cases.

costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.
costl_0	Local cost for medicine and surgery
costl_1	Local cost for rehabilitation
costl_2	Local cost for blind rehabilitation
costl_3	Local cost for spinal cord injury
costl_4	Does not exist; this category is included with 0
costl_5	Local cost for psychiatry
costl_6	Local cost for substance use treatment
costl_7	Local cost for intermediate medicine
costl_8	Local cost for domiciliary
costl_9	Local cost for nursing home care
costl_10	Local cost for psychosocial residential rehabilitation treatment programs
costn*	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers. Same categories as local costs.
los*	Length of stay overall and for the different categories of care. Same categories as local cost.
flag	An indicator for local costs that deviate +/- 2 standard deviations from the national costs.
flagnh	A flag for community nursing home. HERC does not estimate these costs.
flagext	A flag to identify cases where the costs were recalculated because HERC length of stay differed from PTF main length of stay.

A single discharge record provides important subtotals. For example, if a researcher is interested in mental health costs, he/she can now identify the mental health costs for every inpatient encounter. This is particularly helpful for those patients who receive care in many different categories during a stay. Again, note that these changes only pertain to the inpatient discharge datasets.

8.3.2 Acute medical-surgical dataset

This dataset is best described as a discharge dataset for persons who were discharged or transferred from an acute medical-surgical bedsection in the fiscal year. The key to understanding this dataset is that we aggregated the bedsection files to make a discharge file that is analogous to the MedPar dataset (see section 7.1).

The first step of the process involved identifying acute medical-surgical bedsections.[8] If, in a stay,[9] a person was in three acute medical-surgical bedsections, we combined these bedsections.

[8] The medical-surgical bedsections in FY 98 were 01, 02, 03, 04, 05, 06, 07, 08, 09, 10, 11, 12, 14, 15, 16, 17, 18, 19, 24, 31, 34, 35, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 65, 75, 83. All others were considered were considered non medical-surgical.
[9] Stays were defined by five variables: scrssn, sta3n, admitday, adtime, disday.

Transfers within acute medical-surgical bedsections, such as from surgery to medicine, were aggregated into a single record. We adopted the rule that if a patient was transferred from an acute medical-surgical bedsection to another acute medical-surgical bedsection that this would be considered part of the same acute medical-surgical stay. Similarly, if a person was transferred from an acute medical-surgical bedsection to a non-medical/surgical bedsection, we ruled that the acute medical-surgical stay had ended. Transfers from an acute medical-surgical bedsection to a non-medical/surgical bedsection and back to an acute medical-surgical bedsection were treated as one non-medical/surgical and two acute medical-surgical stays.

You will want to link this file to the PTF bedsection files. But before you merge those files with this cost file, you will need to aggregate the bedsection file. We have provided the code for this in Appendix D. You can also contact HERC if you would like an electronic version of this SAS code.

Variables in the Medical Surgical Dataset

scrssn	Numeric field. Identifies a patient's scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center's station number. These can change when facilities merge.
adtime	Admission time for an inpatient stay.
admitday	Admission day for an inpatient stay (SAS date)
disday	Discharge day for an inpatient stay (SAS date).
bsoutday	Discharge day for the bedsection
bsinday	Does not exist; creating the dataset alters this variable. If you really need it, consider making a pseudo-bsinday by subtracting LOS from the bsoutday. However, this may be imperfect for merging.
Source*	numeric field that identifies the source of the data.
	1=XB census
	2=XB discharge
	3=PB census
	4=PB discharge
	5=OBS discharge
	6=OBS census
Los	Length of stay.
Drgwt	Diagnostic related weight created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Numeric field.
Icudays	Length of stay in the ICU; 0 if none.
Drg	Diagnostic related group created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Each group has an associated drgwt– see above. Numeric field.
costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.

Costn	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers.
Flag	An indicator for local costs that deviate +/- 2 standard deviations from the national

*Not included after FY04.

8.3.3 Rehabilitation, mental health or long-term dataset

This dataset contains costs for people who were in non medical-surgical bedsections. Only costs for stays during the fiscal year were included. If a person was admitted and discharged in FY98, then the total cost of their stay is in this FY98 dataset. However, if a person was admitted prior to FY98 (10/1/97), then only the costs for the portion of the stay during FY98 is reported in the dataset. One of the reasons for doing this is that there are some people in long-term care who have been there for 30+ years. It would be extremely difficult to identify the entire cost of these stays. For information on costs prior to FY98, see HERC working paper
(P. G. Barnett, S Chen, & T. H. Wagner, 2000).

Eventually a rehabilitation, mental health or long-term care stay is discharged. Any costs during the year of discharge is captured in this dataset. The total cost of the discharge would then be captured by the discharge dataset; this is calculated by summing together the nominal costs for each fiscal year.

Rehabilitation, Mental Health and Long-Term Care Dataset

scrssn	Numeric field. Identifies a patient’s scrambled social security number
sta3n	3-digit numeric field. Represents the VA medical center’s station number. These can change when facilities merge.
adtime	Admission time for an inpatient stay.
admitday	Admission day for an inpatient stay (SAS date)
disday	Discharge day for an inpatient stay (SAS date).
bsoutday	Discharge day for the bedsection
bsinday	Admit day for the bedsection
bedsection	Lists the bedsection of the treating physician. For more information see http://www.virec.research.med.va.gov/References/RUG/RUG-Inpatient02.pdf
lsb	Length of stay in bedsection.
distype	Type of discharge; identifies death in hospital. See http://www.virec.research.med.va.gov/References/RUG/RUG-Inpatient02.pdf for more information.

source	numeric field that identifies the source of the data. 1=XB census 2=XB discharge 3=PB census 4=PB discharge 5=OBS discharge 6=OBS census
cat	HERC category of care 0= Medicine and Surgery 1= Rehabilitation 2= Blind rehabilitation 3= Spinal cord injury 4= Surgery (category does not exist; we combined it with 0) 5= Psychiatry 6= Substance use treatment 7= Intermediate medicine 8= Domiciliary 9= Nursing Home 10= Psychosocial residential rehabilitation programs
drg	Diagnostic related group created by Centers for Medicare and Medicaid Services for reimbursing inpatient Medicare stays. Each group has an associated drgwt– see above. Numeric field.
costl	Total local cost. Represents the entire cost of the stay, reconciled with the local medical center's expenditures.
costn	Total national cost. Represents the entire cost of the stay, reconciled with expenditures from all VA medical centers.
flag	An indicator local costs that deviate +/- 2 standard deviations from the national costs.

8.3.4 Flag

An important variable is the flag variable. This variable indicates when the local cost estimate (costl) is > 2 standard deviations above or below the national cost estimate. Flag is an indicator or dummy variable; use the costl with caution when the flag variable is one.

8.4 When not to use the average cost dataset

8.4.1 Effects not detected in this cost estimate

It is not always appropriate to use these average cost data in your analysis. The average cost method assigns the same cost to all inpatient stays with the same demographic and discharge information. Stays that have the identical characteristics will have the same cost. If you are interested in assessing the cost consequences of a new procedure, then these data are likely to be inappropriate unless the cost of the procedure is entirely reflected by variables in the cost function (see page 36). If the procedure saves money, but it does not affect one of the variables in the cost function, such as DRG weight or length of stay, then these stays will all get the average cost.

For example, let us assume that we had a new procedure for transfusing blood during a heart transplant. We are interested in whether this new procedure saves money. First, let us assume that this intervention would not affect the patient's DRG. In this case, it is also likely that the intervention would not affect other variables in the cost function, such as length of stay. Therefore, the estimated cost of care for people who received this new procedure would be the same estimated cost of care for people receiving the usual therapy. This does not mean that there was not a cost difference from this new therapy. It only means that any differences were not reflected in the HERC average cost dataset.

8.4.2 Comparison of medical center efficiency

The economic definition of efficiency is to use fewer inputs to make the same level of output, or conversely to use the same number of inputs to make more output. These costs estimates are relative value weights based on Medicare patient discharge characteristics. The local cost estimate is generated by reconciling the relative value weights to the CDR. But, the relative value weights DO NOT capture differences in the quantity or price of the inputs. In addition, the CDR costs (FY98-FY03) and DSS costs (FY04) exclude the cost capital financing. Finally, we distribute other short-term fixed costs in proportion to the variable costs. Although these issues may not be critical for cost-effectiveness analysis, they are more problematic and potentially fatal for efficiency analysis.

8.4.3 Point estimates versus variance estimates

We believe the average cost method produces relatively accurate point estimates for the costs. However, a consequence of estimating costs with a cost function is that the variance of the estimated costs is biased downwards. The reason for this is that many factors that affect costs are not included in the cost function, and if the stays are identical on all observed factors then these cases receive the same estimated cost. In Table 8.4 we show the costs reported by Medicare (1996) for five DRGs. We also show the estimated costs from our cost function (estcost). As is clear from this table, the standard deviation is smaller in the estimated costs. Also, the minimum and maximum are attenuated toward the mean.

Table 8.4: The cost function's effect on the variation of the estimated costs

	Obs	Mean	Std. Dev	Min	Max
DRG14 Specific cerebrovascular disorders except TIA					
cost	10534	6829	7587	7	175346
estcost	10534	7377	7476	685	147135
DRG79 Respiratory infections & inflammations age >17 w cc					
cost	7767	7923	8445	16	213967
estcost	7767	8210	6423	685	198091
DRG88 Chronic obstructive pulmonary disease					
cost	15428	4786	5525	5	203877
estcost	15428	4535	4269	685	128695
DRG89 Simple pneumonia & pleurisy age >17 w cc					
cost	12905	5468	8863	8	662916
estcost	12905	5238	4675	685	160280
DRG127 Heart failure & shock					
cost	21463	4941	4979	10	109945
estcost	21463	5224	4479	685	190673

Note: cost is cost adjusted charges and estcost is the estimated cost adjusted charges.

If you are interested in evaluating the variation of these cost estimates, then use these costs carefully. If you use these cost estimates in a statistical model, most statistical tests will be biased toward the null. If you are trying to identify cost outliers (high or low), then you will almost certainly miss some.

8.5 Duplicates

Researchers who want to merge VA utilization data to our average cost estimates need to be aware that the PTF files have duplicates. There are duplicates within each file (e.g., PB discharge file) and between files (e.g., PB discharge file and XB discharge file). We have removed all duplicates in the average cost datasets before we calculated the costs. To prevent a one-to-many merge, you should delete duplicates from any Austin data that you are working with. The best way to handle this is to run the following command in SAS, which will remove any duplicates with the same information. Note that these commands only identify records that have duplicate values of the sort variables. The records may differ in other respects.

In the acute9x, nacute9x, and dischg9x files, we used:

```
proc sort data=<indata> out=<outdata> nodupkey;
  by scrssn admitday adtime disday sta3n bsinday bsoutday;
```

References

Barnett, P. G. (1997). Research without billing data. Econometric estimation of patient-specific costs. *Med Care, 35*(6), 553-563.

Barnett, P. G., Chen, S., & Wagner, T. H. (2000). Determining the cost of VA care with the average cost method for the 1993-1997 fiscal years. *HERC Technical Report #2*.

Appendix A
Reconciliations for FY98-05

		New category							
sta3n	old_cat	1998	1999	2000	2001	2002	2003	2004	2005
402	6	5	5	5	5	5	5		
405	3				7				
436	7		9						
437	6				5				
438	6	5	5	5	5				
442	5			7					
452	1	7							
452	3			7					
452	6	5							
459	0				9				
459	4				9				
459	7	9							
463	0				8	8		8	
463	9						8		
500	2	1							
500	7	9							
503	4						0		0
504	6	1							
504	7		9						
506	1		9						
508	6	5	5	5	5	5	5		
508	7			9					
509	6		5	5	5				
512	1		9	9	9	9			
515	1		9						
515	6	5	5						
515	8	9	9				9		
516	1	9	9	9	9	9			
516	2		9						
516	6	5	5	5	5	5			
518	0					8			9
518	6					5			
520	6					5			
521	3		2	2					
521	9						2		
523	7	5							
526	1		9	9	9	9	9		
526	6	5							5
528	1		9						
528	3			1					
528	6			5		5	5		

		New category							
sta3n	old_cat	1998	1999	2000	2001	2002	2003	2004	2005
529	4		9						
529	7	9							
531	3			9					
531	6			5					
534	9	0							
537	2	1							
537	7					5	5		
537	8								1
538	6	5							
539	1		9						
540	9						7		
541	7						9		
542	6				5				
543	7		9						
544	7					9		9	
546	2					1			
546	6							5	
549	1				9				
549	6	5	5	5					
550	6					5	5		
552	1						9		
552	6						5		
552	7						9		
553	1		9						
554	7		9	9	9	9	9		
555	1	9	9						
555	6	5	5						
555	10	5							
556	4							0	0
556	6	5	5	5	5	5			
556	7					9			
557	1		9	9					
557	6				8	8			
558	6	5							
558	7	9	9	9	9	9	9		
561	2	1							
561	6		5	5	5		5		
562	5	7							
567	0				9				
570	6				5	5	5		
570	7					9	9		
573	6	5							
573	10	5							
578	7					9			
580	6						5		

		New category							
sta3n	old_cat	1998	1999	2000	2001	2002	2003	2004	2005
581	5	7							
585	6			5	5	5	5		
586	6					5	5		
586	7						9		
589	6	5	5			5			
589	7	5	9	5					
590	6				5				
596	1			9					
597	6		5						
598	6	5			5				
600	6						5		
603	1	7							
603	3	7							
603	9						7		
605	1	9	9	9	9	9	9		
608	1			9	9				
608	4			0	0	0	0		
608	7						9		
609	7	9							
610	1		9	9					
610	3		0						
610	4			0	0	0	0		
610	7	9	9						
612	1			9					
612	5			9	9				
612	7	9	9	9					
614	1	7					7		
614	9					7	7		
619	6	5	5						
619	8	9	9	9					
620	4	0							
620	6				5				
620	10				5				
621	6	5	5	5	5	5	5		
621	7			9					
622	6	5							
623	1		7						
623	9						7		
626	6					5	5		
629	8				9				
630	6			5	5				
631	6	5	5	5					
631	7					9			
631	8					9			
632	2	1				9	9		

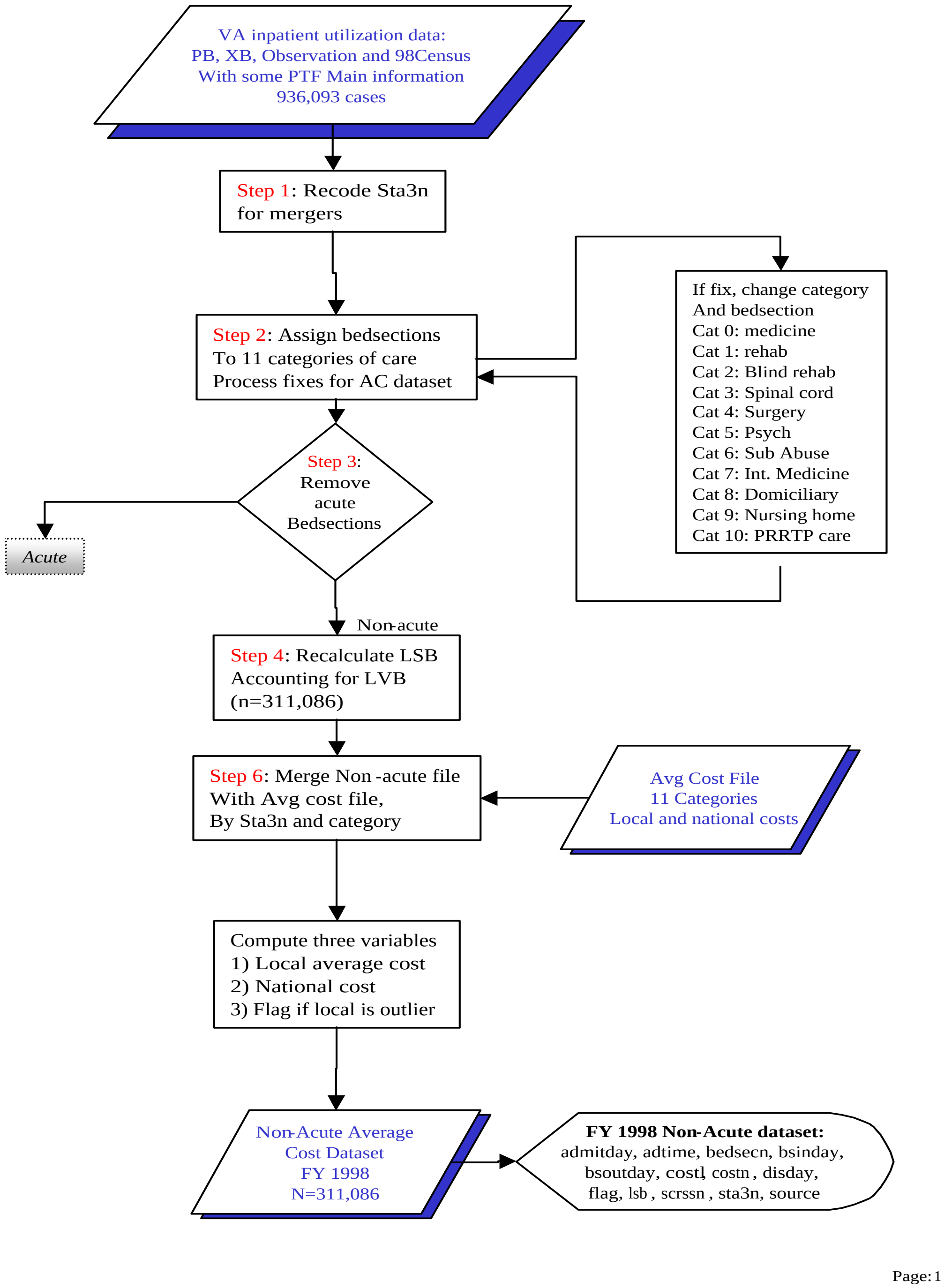
		New category							
sta3n	old_cat	1998	1999	2000	2001	2002	2003	2004	2005
632	6				5	5			
635	6			5					
636	1			9	9	9	9		
637	6					5			
642	1		9	9					
642	6	5	5	5		5	5		
642	7	9				9			
644	1	9	9	9	9				
644	3				9				
644	6					5			
646	6	5							
647	5	9							
648	1						9		
648	8						9		
649	1							9	
652	6			5					
653	6		5	5	5	5	5		
654	1			9	9	9	9		
654	6			5	5	5	5		
655	3		0						
655	4	0		0	0	0	0		
656	6			5					
656	7				9				
657	1							9	9
657	6				5				
658	6		5						
660	9					7	7		
662	2						9		
662	3					9			
662	7					9	9		
662	8					9		9	
662	10						5		
664	1	9			9				
664	6		5	5	5	5	5		
667	7	5							
668	1			9	9				
668	6	5							
670	3		1						
671	1		9	9					9
671	7		9	9	9	9			
672	7						9		
673	6					5			
674	1	9	9	9	9	9	9		
674	6	5	5	5	5				
678	1				9				

		New category							
sta3n	old_cat	1998	1999	2000	2001	2002	2003	2004	2005
678	6	5	5	5	5	5	5		
678	7		9	9		9			
679	0				7	7			
679	1				9				
679	8	9	9						
687	4	0							
688	6	5	5			5	5		
689	3						7		7
689	7	9	9	9					
689	8						9		
691	3	1							
691	6	5	5						
692	0	8	8	8	8	8			
692	9						8		
693	6					5	5		
695	7					9			

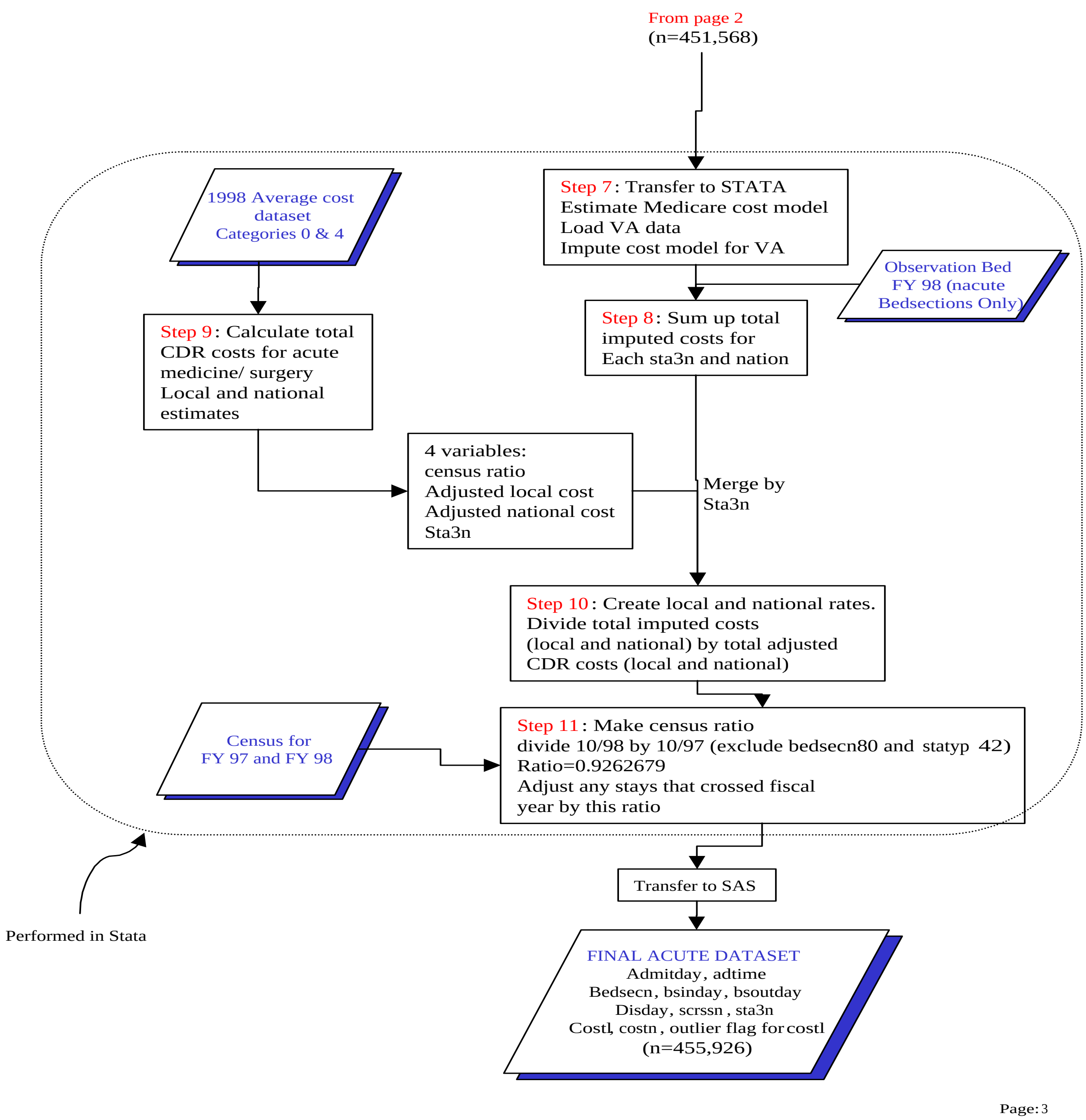
Note: if the cell is blank for a new category year, then there were no reconciliations made

Appendix B
Flow diagram for inpatient care

Development of non-acute average cost dataset



Scaling the average cost dataset for acute inpatient costs in FY1998



Appendix C
VHA directive on observation beds

Department of Veterans Affairs **VHA DIRECTIVE 98-025**
Veterans Health Administration
Washington, DC 20420 **May 5, 1998**

RECORDING OBSERVATION AND/OR SHORT-STAY PATIENTS
1. PURPOSE: This Veterans Health Administration (VHA) Directive provides VHA policy for the definition and recording of observation and/or short-stay patients.

2. BACKGROUND: As outlined in the "Vision for Change," VHA will place patients in the most appropriate setting. In many instances, this involves "observing" a patient for an extended period of time without admitting them as an inpatient. While observation units are considered to be outpatient or ambulatory services, current software supporting Nutrition and Food and Pharmacy Services only work for inpatient beds. Properly recording the level of services while maintaining automated support for functional activities will require a creative approach to classifying services to these patients. This policy also complies with current Health Care Financing Association (HCFA) guidelines used in the administration of the Medicare program.

3. DEFINITION

a. Observation Patient. An observation patient is one who presents with a medical condition with a significant degree of instability or disability, and who needs to be monitored, evaluated and assessed for either admission to inpatient status or assignment to care in another setting. An observation patient can occupy a special bed set aside for this purpose or may occupy a bed in any unit of a hospital, i.e., urgent care, medical unit. These types of patients should be evaluated against standard inpatient criteria. These beds are not designed to be a holding area for Emergency Rooms. The length-of-stay in observation beds will not exceed 23 hours.

b. Lodger. A lodger is not an observation patient. By definition a lodger does not receive healthcare services.

NOTE: Routine post-procedure recovery from ambulatory surgery is not observation. Examples: Recovery from a cardiac catheterization and release from the facility within 6 hours of the completion of the catheterization would not constitute post-surgical observation since the normal recovery time is 4 to 6 hours. A patient may report to the medical center for laser removal of cataracts. During the laser procedure, the patient may have a reaction to some of the medication and would be admitted to the appropriate bed section for evaluation of the reaction.

4. POLICY: To accomplish this policy within the context of VHA's supporting software, patients will be assigned to a treating specialty code of Observation. All services and costs associated with Observation treating specialties will be captured and assigned to inpatient services.

THIS VHA DIRECTIVE EXPIRES MAY 5, 2003
5. ACTION

a. The following Patient Treatment File (PTF) Treating Specialties and Cost Distribution Report (CDR) account numbers are to be utilized for recording Observation patient activity.

Treating Specialty	PTF #	CDR #
Medical Observation	24	1110.00
Surgical Observation	65	1210.00
Psychiatric Observation	94	1310.00
Neurology Observation	18	1111.00
Blind Rehabilitation Observation	36	1115.00
Spinal Cord Injury Observation	23	1116.00
Rehabilitation Medicine Observation	41	1113.00

b. These Treating Specialties should be utilized when setting up Observation Units. The following guidelines and menu options will assist you. Using the Ward Definition menu option create Observation Unit wards.

The Treating Specialty should be one of the above Observation Treating Specialties appropriate for the ward location. The service for the Observation Unit ward should be NON-COUNT. Remember to include the Gain and Losses Sheet (G&L) location. Using the Treating Specialty Set-up option, set up the new Treating Specialties.

c. Patients placed on Observation status will be admitted to one of the treating specialties listed above. This will enable the facility to track the patients on the G & L, and use the required Pharmacy and Nutrition and Food Services software to deliver services. An observation patient requiring subsequent admission would be released from Observation status by discharging them from the facility and then admitting them to an acute care-treating specialty.

d. Patients already designated as inpatient status must be discharged and re-admitted to an Observation Treating Specialty for no more than the time limits previously indicated (especially normal ambulatory surgery which are not related to the reason for hospitalization). Following the Observation period, the patient must be re-admitted to inpatient status, if further hospitalization is required. Nursing Home care Unit (NHCU) and Domiciliary (DOM) patients requiring Observation services would be transferred Absent Sick in Hospital (ASIH) from the NHCU or DOM and admitted to an Observation Treating Specialty.

e. Insurance carriers of patients on Observation status will be billed at the appropriate inpatient rate for the medical, surgical or psychiatric bed section using revenue code 760, until such time as an observation unit rate can be established. This is a facility charge and should be billed on an Uniform Billing Form (UB)-92. For billing professional fees only, Current procedural Terminology (CPT) codes should be used. A principal diagnosis should be available for these patients at the time the patient is either discharged and re-admitted to another treating specialty for inpatient care or to an appropriate ambulatory care setting.

f. First party patient charges for Category C observation patients will be billed at the published Category C outpatient visit copayment rate.

g. Utilizing this data report methodology will enable data users to separate the activity of these patients for their purposes. For performance measurement purposes, these patients would NOT be included as acute care inpatients. Procedures performed while a patient is assigned to Observation status will be considered ambulatory for performance measure purposes.

h. Facilities will complete and transmit PTF records for reporting Observation patients when discharged from Observation status. If a patient were admitted following observation, the acute care PTF record would be transmitted after discharge from inpatient care. Attachment A outlines the minimal requirements for patient record documentation of Observation patients.

I. Facilities will complete and transmit PTF records for reporting Observation patients when discharged from Observation status. If a patient were admitted following observation, the acute care PTF record would be transmitted after discharge from inpatient care. Attachment A outlines the minimal requirements for patient record documentation of Observation patients.

j. Patch DG*5.3*176 is being released to implement this directive. Appropriate IB patches will be released in the future.

6. REFERENCES:

Glossary of Healthcare Terms, American Health Information Management Association, 1994, page 14.

7. FOLLOW-UP RESPONSIBILITY

a. For issues affecting classification of patients, Health Administration Service (10C3). Questions concerning classification may be addressed to Kay Evans at (202) 273-8306.

b. For issues concerning billing, Medical Care Cost Recovery (174), Questions concerning billing may be addressed to Nancy Howard at (202) 273-8198.

8. RESCISSIONS: This VHA Directive will expire May 5, 2003.

S/ Thomas Garthwite, M.D.for
Kenneth W. Kizer, M.D., M.P.H.
Under Secretary for Health
Attachment

DISTRIBUTION:	CO:	E-mailed 5/5/98
	FLD:	RD, MA, DO, OC, OCRO and 200 - FAX 5/5/98
	EX:	Boxes 104, 88, 63, 60, 54, 52, 47 and 44 - FAX 5/5/98

Attachment A

Observation Patient Record

Documentation Requirements

Document/ Item	Completion Time	Components of Document Required
Admission Order	On Admission	Timed and dated order for admission of the patient to an Observation Bed
Initial Assessment and History and Physical (H&P)	Immediately	Initial Assessment and screening of physical, psychological (mental) and social status to determine the reason why the patient is being admitted to an Observation Bed, the type of care or treatment to be provided, and the need for further assessment. An extensive Emergency Room (ER) note or Progress Note, documented by the admitting physician, which encompasses the normal criteria for an H&P will suffice as an initial assessment and H&P for the Observation patient.
Progress Notes	Within 8 hours - with subsequent notes documented as the patient's condition warrants. 24 hour re-assessments should be documented	Progress Notes should reflect the status of the patient's condition, the course of treatment, the patient's response to treatment and any other significant findings apparent at the time the progress note is documented. Reassessments should include a plan for (1) discharge or transfer; (2) readmission to inpatient status; or (3) continued observation with evaluation and rationale.
Discharge Order	On Discharge	Timed and dated order for discharge from the Observation status.
Discharge Diagnoses	On Discharge	Complete listing of all final diagnoses including complications and comorbidities.
Discharge Note	On Discharge	Summarization of the reason for the Observation admission, the outcome, follow-up plans and patient disposition, and discharge instructions (diet, activity, medications, special instructions). *NOTE: This document may be written in the Progress Notes or dictated, according to local policy.*

Appendix D
Contents of HERC DATASET at Austin

Discharge dataset

```
             -----Alphabetic List of Variables and Attributes-----

 # Variable Type Len Pos Format  Informat Label
--------------------------------------------------------------------------------
 4 ADMITDAY Num    5 305 DATE9.  7.       DATE OF ADMISSION (SASDATE)
 3 ADTIME   Num    5 300                  TIME OF ADMISSION
 6 B4FY98   Num    8   0                  FLAG if Admitted Prior to Fiscal Year 98
10 COSTL_0  Num    8  32                  cost (local) for cat 0:acute med/surg
11 COSTL_1  Num    8  40                  cost (local) for cat 1:rehab
12 COSTL_2  Num    8  48                  cost (local) for cat 2:blind rehab
13 COSTL_3  Num    8  56                  cost (local) for cat 3:spinal cord
14 COSTL_5  Num    8  64                  cost (local) for cat 5:psych
15 COSTL_6  Num    8  72                  cost (local) for cat 6:substance abuse
16 COSTL_7  Num    8  80                  cost (local) for cat 7:intermed. med
17 COSTL_8  Num    8  88                  cost (local) for cat 8:domiciliary
18 COSTL_9  Num    8  96                  cost (local) for cat 9:nursing home
19 COSTL_10 Num    8 104                  cost (local) for cat 10:PRRTP
30 COSTN_0  Num    8 192                  cost (national) for cat 0:acute med/surg
31 COSTN_1  Num    8 200                  cost (national) for cat 1:rehab
32 COSTN_2  Num    8 208                  cost (national) for cat 2:blind rehab
33 COSTN_3  Num    8 216                  cost (national) for cat 3:spinal cord
34 COSTN_5  Num    8 224                  cost (national) for cat 5:psych
35 COSTN_6  Num    8 232                  cost (national) for cat 6:subst. abuse
36 COSTN_7  Num    8 240                  cost (national) for cat 7: intermed. med
37 COSTN_8  Num    8 248                  cost (national) for cat 8: domiciliary
38 COSTN_9  Num    8 256                  cost (national) for cat 9: nursing home
39 COSTN_10 Num    8 264                  cost (national) for cat 10: PRRTP
 5 DISDAY   Num    5 310 DATE9.  7.       DATE OF DISCHARGE (SASDATE)
41 FLAGEXT  Num    8 280                  FLAG if Observation Days/Cost Extrapolated
40 FLAGNH   Num    8 272                  Community Nursing Home Discharge
20 LOS_0    Num    8 112                  length of stay for cat 0:acute med/surg
21 LOS_1    Num    8 120                  length of stay for cat 1:rehab
22 LOS_2    Num    8 128                  length of stay for cat 2:blind rehab
23 LOS_3    Num    8 136                  length of stay for cat 3:spinal cord
24 LOS_5    Num    8 144                  length of stay for cat 5:psych
25 LOS_6    Num    8 152                  length of stay for cat 6:substance abuse
26 LOS_7    Num    8 160                  length of stay for cat 7:intermed. med
27 LOS_8    Num    8 168                  length of stay for cat 8:domiciliary
28 LOS_9    Num    8 176                  length of stay for cat 9:nursing home
29 LOS_10   Num    8 184                  length of stay for cat 10:PRRTP
 1 SCRSSN   Num    7 288 SSN11.  11.      SCRAMBLED SOCIAL SECURITY NUMBER
 2 STA3N    Num    5 295 STA3NL.          STATION (PARENT)
 8 costl    Num    8  16                  case-mix adj local cost
 7 costn    Num    8   8                  case-mix adj national cost
 9 flag     Num    8  24                  Cost Estimate +/- 2 Std. from Average


Sortedby:      SCRSSN ADMITDAY ADTIME DISDAY STA3N
```

Rehabilitation, Mental Health and Long Term Care Dataset

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Label
1	ADMITDAY	Num	8	0	MMDDYY10.	DATE OF ADMISSION (SASDATE)
7	ADTIME	Num	8	48		TIME OF ADMISSION
8	BEDSECN	Num	8	56	BEDSECN.	BED SECTION
3	BSINDAY	Num	8	16	MMDDYY10.	DAY ADMITTED TO BEDSECT (SASDATE)
4	BSOUTDAY	Num	8	24	MMDDYY10.	DAY TRANSFERED FROM BEDSECT (SASDATE)
2	DISDAY	Num	8	8	MMDDYY10.	DATE OF DISCHARGE (SASDATE)
10	DISTYPE	Num	8	72	DISTYPEL.	TYPE OF DISCHARGE
9	LSB	Num	8	64		LENGTH OF STAY IN BEDSECTION
5	SCRSSN	Num	8	32	SSN11.	SCRAMBLED SOCIAL SECURITY NUMBER
6	STA3N	Num	8	40	STA3NL.	STATION (PARENT)
12	cat	Num	8	88		Category of Care
14	costl	Num	8	104		Local-level Cost Estimate
15	costn	Num	8	112		National-level Cost Estimate
13	flag	Num	8	96		Cost Estimate +/- 2 Std. from Average
11	source	Num	8	80		Categorical Indicator of Type Bedsection File Input

Sortedby: SCRSSN ADMITDAY ADTIME BSINDAY BSOUTDAY DISDAY STA3N

Medical Surgical Care Dataset

-----Alphabetic List of Variables and Attributes-----

#	Variable	Type	Len	Pos	Format	Label
4	ADMITDAY	Num	8	24	MMDDYY10.	DATE OF ADMISSION (SASDATE)
3	ADTIME	Num	8	16		TIME OF ADMISSION
6	BSOUTDAY	Num	8	40	MMDDYY10.	DAY TRANSFERED FROM BEDSECT (SASDATE)
12	COSTL	Num	8	80		Local-level Cost
13	COSTN	Num	8	88		National-level Cost
5	DISDAY	Num	8	32	MMDDYY10.	DATE OF DISCHARGE (SASDATE)
11	DRG	Num	8	72		Diagnostic Relate Groupings(DRG)
14	FLAG	Num	8	96		Cost Estimate +/- 2 Std. from Average
1	SCRSSN	Num	8	0	SSN11.	SCRAMBLED SOCIAL SECURITY NUMBER
7	SOURCE	Num	8	48		Categorical Indicator of Type Bedsection File Input
2	STA3N	Num	8	8	STA3NL.	STATION (PARENT)
9	drgwt	Num	8	56		Diagnostic Related Groupings(DRG) Weights
10	icudays	Num	8	64		Number of days in an Intensive Care Unit
8	los	Num	5	104		LENGTH OF STAY IN BEDSECTION

Sortedby: SCRSSN ADMITDAY ADTIME BSOUTDAY DISDAY STA3N